COVARIANCE STRUCTURE ANALYSIS
WITH
POLYTOMOUS AND INTERVAL DATA

by

Yin-Ping Leung

A

Thesis

Submitted to

(Division of Statistics)

The Graduate School

of

The Chinese University of Hong Kong

In Partial Fulfillment

of the Requirements of the Degree of

Master of Philosophy

(M. Phil.)

May, 1992

THE CHINESE UNIVERSITY OF HONG KONG

GRADUATE SCHOOL

The undersigned certify that we have read a thesis, entitled "Covariance Structure Analysis with Polytomous and Interval Data" submitted to the Graduate School by Yin-Ping Leung (梁燕萍) in partial fulfillment of the requirement for the degree of Master of Philosophy in Statistics. We recommend that it be accepted.

W. Y. Poon.

Dr. W. Y. Poon,

Supervisor.

Dr. S. Y. Lee

Dr. K. H. Wu

Prof. P. M. Bentler,

External Examiner.

DECLARATION

No portion of the work referred to in this thesis has been submitted in support of an application for another degree or qualification of this or any other university or other institution of learning.

ACKNOWLEDGEMENT

I would like to express my deep gratitude to my supervisor, Dr. W. Y. Poon for her everlasting encouragement, invaluable advices and guidance during the course of this research program.

Moreover, I would like to extend my gratitude to the entire staff of the Department of Statistics, especially to Mr. M. L. Tang and Mr. K. H. Leung for their kind assistance.

Lastly, I have to thank my family for their everlasting support.

ABSTRACT

The analysis for the structural equation models with polytomous and interval data are discussed. In the first part, the partition maximum likelihood estimate for correlations among the polytomous and interval variables is developed. The behavior of the estimate is investigated by a simulation study. Then the analysis of the structural equation models is considered in the second part. By using different identification methods to identify the parameters corresponding to the polytomous variables, two types of the structural equation models are possible, the covariance and correlation structure models. Hence, two different approaches to handle these models are developed. A three-stage procedure is established for the covariance structure model. The pseudo partition maximum likelihood method is used. In the first stage, the partition maximum likelihood estimates of the thresholds are obtained. The second stage estimates the elements of the covariance matrix via the pseudo maximum likelihood method. It will be shown that the asymptotic distribution of these estimates are jointly multivariate normal. The estimates of structural parameters are obtained by generalized least squares approach with a correctly specified weight matrix in the third stage. Asymptotic properties of the structural parameter estimate will also be provided. A Monte Carlo study is conducted to investigate the performance of this estimate.

A two-stage procedure is developed for the correlation structure model. The partition maximum likelihood estimates of the elements in the correlation matrix are obtained in the first stage. From these estimates

and their asymptotic distributions, estimates of the structural parameters are obtained via the generalized least squares method in the second stage. Basic statistical properties of these estimates are derived and a Monte Carlo study is also presented. Moreover, some examples are presented to compare the accuracy of the two methods with same underlying correlation structure.

CONTENTS

Chapter 1

Introduction

Structural equation modeling is a statistical technique in analyzing the structural parameter $\underline{\theta}$ of a covariance structure model, $\underline{\Sigma}(\underline{\theta})$. This method has drawn much attention in applied statistics and now has been well accepted as one of the most important statistical tools for analyzing behavioral and social data. Traditionally, this multivariate technique is used to handle data which are observable and continuous. However, in many real-life situations, these assumptions are violated due to various reasons such as the design of questionnaire or the uncertainty of the subjects about the variable under study. Therefore, it is important to develop statistical techniques to cope with the data with various kinds of fuzziness.

When continuous latent variables are only observable in categorical form, they are called polytomous variables. In practice, especially in behavioral and social studies, most data are in dichotomous or polytomous forms. Examples of these polytomous variables are attitude items, performance items and the like. A typical case is that a patient with fever is asked to describe his feeling about the statement "extent of pain relief after treatment" as (1) Worse, (2) Same, (3) Slight improvement, and (4) Marked improvement. It is an example in which a continuous variable underlies a polytomous observed variable. When analyzing this kind of data, direct applying of the standard statistical method with the continuous assumption may result a misleading conclusion (see, e.g. Olsson, 1979a). Hence, the analysis of polytomous data has recently received a great attention in the literature. Statistical methods based on different assumptions in analyzing this kind of data have been developed. In particular, Lee, Poon, & Bentler (1990a) developed a full maximum likelihood method for covariance structure analysis with polytomous variables. Later, a three-stage more efficient procedure for the analysis has been established by Lee, Poon, & Bentler (1990b).

Besides the polytomous data, there are also data with another form of fuzziness, namely, the interval data, see, for example, Poon, Lee, & Bentler (1991). Instead of giving a precise point observation, all the random observations given are intervals. The interval data can be viewed as the data in between the polytomous and continuous data. They do not describe the system as precise as the continuous data but carries more information than the polytomous data.

In collecting polytomous data, subjects are forced to choose one from a small number of categories with unknown thresholds. That is, every subject is assumed to share the same thresholds in defining his intervals. However, the subject may have his own thresholds. For example, a student may consider that he has done a good job if he gets a grade point average (GPA) over 3.0, but another student may consider that he does not satisfy with his performance if he gets the same GPA. Moreover, in some cases, the subject may find it difficult to precisely describe his feeling if the number of categories is too small, or equivalently, the interval length is too large. With the "extent of pain relief after treatment" question as an example, a patient is asked to choose one of the four categories. He may hesitate to respond if he feels complete relief of the pain but not "Marked improvement", but he has to choose the same category " Marked improvement" .

in both cases. Therefore, some information is surely lost. Hence, in these cases, it is recommended to use the interval data.

In obtaining interval data, subjects are asked to provide an interval instead of choosing a category or giving a precise point value. For example, in collecting information about the question "extent of pain relief after treatment", the subject is asked to provide an interval within a pre-assigned range, say (-10,10). The value -10 stands for the worsening of the pain while 10 stands for the complete relief of the pain. Thus, the subject has the freedom to choose his own intervals within the provided range. Moreover, he can also determine the length of the intervals. For example, he can give a large interval if he is uncertain about his feeling. Poon, Lee, & Bentler (1991) have investigated the estimation of correlation of two variables with interval data. The statistical theories based on the partition maximum likelihood method for covariance structure analysis are also established (Poon & Lee, 1991).

In practice, interval and polytomous data may be encountered simultaneously, for instance, observations of effect of a drug on patients with fever are taken. Each patient is given the same amount of a drug and then asked to respond to two statements, "Do you like taking this drug?" and "extent of pain relief after treatment". He is asked to choose one from the three categories (1) Like (2) Don't know and (3) Dislike for the first statement and to provide an interval to describe his feeling for the second statement. In this case, the response to the first statement is a polytomous observation while the response to the second statement is an interval observation. Although methodologies for analyzing these two kinds of data separately have been well-developed, they cannot be applied in here.

Therefore, methods for analyzing these types of observations simultaneously are strongly desired.

The main purpose of this thesis is to develop basic theories for analysis of structural equations models with interval and polytomous data. The estimates of all the correlations involved are obtained via the partition maximum likelihood estimation and the details will be presented in Chapter 2. Then, a three-stage procedure is employed to obtain the structural parameters of the general structural equation model in Chapter 3. In the first stage, the partition maximum likelihood estimates of the thresholds are obtained. In the second stage, the pseudo partition maximum likelihood method is used to estimate the elements in the covariance matrix, $\underline{\Sigma}$, without imposing any structure and with the thresholds fixed at the values in first stage. The asymptotic distribution of the pseudo estimates are derived and a correctly specified weight matrix is also given in this stage. The third stage estimates the structure via the generalized least squares approach. Asymptotic properties for further statistical inference of the model are given. By another identification method, a two-stage procedure for analyzing the correlation structure model is developed and the materials is presented in Chapter 4. In stage one, the estimates of the correlations are obtained via the partition maximum likelihood estimation. Then, the generalized least squares method is employed to estimate the structural parameters with an appropriate weight matrix given in stage one. Statistical inferences can be made on the model based on the asymptotic distribution of the estimates. Based on the written computer programs for the partition maximum likelihood estimates of the correlations among the two kinds of data, the three-stage procedure for the covariance structure analysis and the two-stage procedure for the correlation structure analysis, Monte Carlo studies are conducted to investigate the behaviors of the estimates discussed. Moreover, several examples are presented to compare the performances of the two methods under the same correlation structure.

Chapter 2

Estimation of the Correlation between Polytomous and Interval Data

As mentioned earlier, an interval variate can be viewed as an polytomous variate without fixing the thresholds. Correlation between these two kinds of variables is of interests. The main objective of this chapter is to obtain an estimate for the parameter vector, which involved the correlation, the mean, and the thresholds of the variables, of the model discussed in Section 1. Two estimation methods are developed and their asymptotic properties are also derived. They are the classical maximum likelihood (ML) procedure and the partition maximum likelihood (PML) procedure which are discussed in Sections 2 and 3 respectively. In Section 4, a Monte Carlo study is conducted to study the performance of the estimator.

§ 2.1 Model

Let $\underline{X}$ and $\underline{Y}$ be two latent continuous random vectors with dimensions p_1 and p_2 respectively. Suppose that $(\underline{X}',\underline{Y}')'$ is coming from a multivariate normal distribution with mean vector $\underline{\mu} = (\underline{\mu}_x', \underline{\mu}_y')'$ and variance covariance matrix $\underline{\Sigma}$ with dimension p x p,

$$\underline{\Sigma} = \begin{bmatrix} \underline{\Sigma}_{xx} & \underline{\Sigma}_{xy} \\ \underline{\Sigma}_{yx} & \underline{\Sigma}_{yy} \end{bmatrix},$$

where $\underline{\mu}_x$ is the $p_1 \times 1$ dimensional mean vector of $\underline{X}$,
$\underline{\mu}_y$ is the $p_2 \times 1$ dimensional mean vector of $\underline{Y}$,
$\underline{\Sigma}_{xx}$ is the $p_1 \times p_1$ covariance matrix of $\underline{X}$,
$\underline{\Sigma}_{yy}$ is the $p_2 \times p_2$ covariance matrix of $\underline{Y}$,
$\underline{\Sigma}_{xy}$ is the $p_1 \times p_2$ covariance matrix of $(\underline{X},\underline{Y})$,

and $p = p_1 + p_2$.

Now, suppose that both $\underline{X}$ and $\underline{Y}$ cannot be observed directly. Instead, we can observe a polytomous random vector $\underline{Z}=(Z_1,\dots,Z_{p_1})'$ and interval random vectors $\underline{W}_1,\dots,\underline{W}_{p_2}$ which are corresponding to $\underline{X}$ and $\underline{Y}$ respectively. The relationship between $\underline{X}$ and $\underline{Z}$ is given by

$$Z_i = k(i) \quad \text{if} \quad \alpha_{i,k(i)} \le X_i < \alpha_{i,k(i)+1}\,, \tag{2.1}$$

for $k(i) = 1,\dots,n(i)$, $i = 1,\dots,p_1$; where $\{\ \alpha_{i,1}= -\infty,\ \alpha_{i,2},\dots,\ \alpha_{i,n(i)},\ \alpha_{i,n(i)+1} = \infty\ \}$ are the thresholds and $n(i)$ is the number of categories corresponding to the ith variable of $\underline{Z}$. It is clear that the variable Z_i takes the values from 1 to $n(i)$.

Unlike a continuous random vector, $\underline{W}_j$ gives interval observations instead of point observations, that is, for each random variable Y_j, $(w_{j,1},w_{j,2})$ is the corresponding mearsurement for $j = 1,\dots,p_2$ and $w_{j,1} < w_{j,2}$. In other words, the examination subject does not certain where he is on a continuous line, rather he can just give out an interval where he belongs to. In this chapter, our main concern is to develop a procedure to estimate the unknown parameter vector $\underline{\beta}$ which contains the thresholds of the polytomous variables, $\underline{\alpha}$, the mean vector, $\underline{\mu}$, and the lower triangular elements in $\underline{\Sigma}$.

Suppose a random sample of polytomous and interval data with size N is given by

$$D_t=\left[\ k_t(1),\dots,k_t(p_1),\underline{w}_{t,1},\dots,\underline{w}_{t,p_2}\right]\,,$$

where t is the subject index running from 1 to N, $k_t(i)$'s are the corresponding polytomous measurements for $i = 1,\ldots,p_1$ and $\underline{w}_{t,j}=(w_{t,j,1}, w_{t,j,2})$ are the corresponding interval measurements for $j = 1,\ldots,p_2$. We also define $\alpha_{t,i,k_t(i)} = \alpha_{i,k(i)}$ for $k_t(i) = k(i)$ and $t = 1,\ldots,N$. The introduction of the subject index t to the measurements of the ith polytomous variable, $k_t(i)$, may be rather misleading as it seems to be in contradiction with the property of the polytomous variable that the thresholds of the variable for every subject are the same. But the use of the subject index t can give a clearer and better presentation when making inference based on both types of the data.

§2.2 Maximum Likelihood Estimation

Let

$$\xi_{k_t(1),\ldots,k_t(p_1),\underline{w}_{t,1},\ldots,\underline{w}_{t,p_2}}$$
$$= \Pr(Z_1=k_t(1),\ldots,Z_{p_1}=k_t(p_1), w_{t,1,1}\le Y_1<w_{t,1,2},\ldots,w_{t,p_2,1}\le Y_{p_2}<w_{t,p_2,2}) \tag{2.2}$$

be the probability that a particular observation for subject t is taken. Since $(\underline{X}',\underline{Y}')'$ is distributed according to $N[\underline{\mu}, \underline{\Sigma}]$, it can be derived from the multivariate normal distribution theory that

$$\xi_{k_t(1),\ldots,k_t(p_1),\underline{w}_{t,1},\ldots,\underline{w}_{t,p_2}} = (-1)^p \sum_{i(1)=0}^{1} \cdots \sum_{i(p)=0}^{1} (-1)^{\sum_{u=1}^{p} i(u)} \times$$

$$\Phi_p(\alpha_{t,1,\nu_t(1)},\ldots,\alpha_{t,p_1,\nu_t(p_1)},w_{t,1,m(p_1+1)},\ldots,w_{t,p_2,m(p)};\underline{\mu},\underline{\Sigma}),$$

(2.3)

where $p = p_1 + p_2$; $\nu_t(j) = k_t(j) + i(j)$ for $j = 1,\ldots,p_1$; $m(s) = 1 + i(s)$ for $s = p_1+ 1,\ldots,p$; and $\Phi_p(a_1,\ldots,a_p; \underline{\mu}, \underline{\Sigma})$ is the distribution function of $N[\underline{\mu},\underline{\Sigma}]$ evaluated at the point $(a_1,\ldots,a_p)$. Hence, $\xi_{k_t(1),\ldots,k_t(p_1),\underline{w}_{t,1},\ldots,\underline{w}_{t,p_2}}$ is a function of the thresholds $\{\alpha_{i,k(i)};i=1,\ldots,p_1;\ k(i) = 2,\ldots,n(i)\}$, the elements of $\underline{\mu}' = (\underline{\mu}_x', \underline{\mu}_y')$ and the lower triangular elements of $\underline{\Sigma}$ $\{\sigma_{ij};i,j = 1,\ldots,p,i\le j\}$.

For the part of the polytomous variables, let us consider the transformed variates $\underline{X}^* = \underline{D}^{-1}(\underline{X}-\underline{\mu}_x) + \underline{\mu}^*$ for any vector $\underline{\mu}^*$ and any diagonal matrix $\underline{D}$ with diagonal elements $d_{ii} > 0$. We have $E(\underline{X}^*) = \underline{\mu}^*$, $\mathrm{cov}(\underline{X}^*) = \underline{\Sigma}^* = \underline{D}^{-1}\underline{\Sigma}_{xx}\underline{D}^{-1}$ and it can be shown that

$$\Phi_{p_1}(\alpha_1,\ldots,\alpha_{p_1};\underline{\mu}_x,\underline{\Sigma}_{xx}) = \Phi_{p_1}(\alpha_1^*,\ldots,\alpha_{p_1}^*;\underline{\mu}^*,\underline{\Sigma}^*)$$

where $\alpha_i^* = d_{ii}^{-1}(\alpha_i - \mu_i) + \mu_i^*$. Hence, the parameters $\underline{\alpha}_j' =(\alpha_{j,2},\ldots,\alpha_{j,n(j)})$, $j = 1,\ldots,p_1$, $\underline{\mu}_x$ and $\underline{\Sigma}_{xx}$ are not identified in the case of polytomous variables. To handle this identification problem, we have to fix diag $\underline{\Sigma}_{xx} = \underline{I}$ and $\underline{\mu}_x = \underline{0}$, thus, we can write the distribution of $\underline{X}$ as $N[\underline{0},\underline{R}_{xx}]$ where $\underline{R}_{xx}$ is the correlation matrix of $\underline{X}$. Consequently, $\underline{\mu} = (\underline{0}', \underline{\mu}_y')'$ and $\underline{\Sigma} = \begin{bmatrix}\underline{R}_{xx} & \underline{C}_{xy}\\ \underline{C}_{yx} & \underline{\Sigma}_{yy}\end{bmatrix}$ where $\underline{C}_{xy}$ is the covariance martix between $\underline{X}$ and $\underline{Y}$. As a result, $\xi_{k_t(1),\ldots,k_t(p_1),\underline{w}_{t,1},\ldots,\underline{w}_{t,p_2}}$ will become a function of the thresholds $\underline{\alpha}$, $\{\alpha_{i,k(i)};\ i = 1,\ldots,p_1;\ k(i) = 2,\ldots,n(i)\}$, the elements of $\underline{\mu}_y$,

$\{\mu_{y,1},\ldots,\mu_{y,p_2}\}$, and the lower triangular elements of $\underline{\Sigma}$, $\underline{\sigma}$, $\{\rho_{ij};\ i,j = 1,\ldots,p_1,\ i>j;\ \sigma_{ij};\ i = 1,\ldots,p_1,\ j = p_1+1,\ldots,p;\ \sigma_{ij};\ i,j = p_1+1,\ldots,p,\ i\geq j\}$. Thus, the parameter vector $\underline{\beta}$ now contains $\underline{\alpha}$, $\underline{\mu}_y$ and $\underline{\sigma}$ as its elements. That is,

$$\underline{\beta}' = (\ \underline{\alpha}'\ ,\ \underline{\mu}_y'\ ,\ \underline{\sigma}'\ ).$$

The likelihood function of the random sample D_t is given by

$$\prod_{t=1}^{N} \xi_{k_t(1),\ldots,k_t(p_1),\underline{w}_{t,1},\ldots,\underline{w}_{t,p_2}}. \tag{2.4}$$

The maximum likelihood (ML) estimate $\hat{\underline{\beta}}$ of $\underline{\beta}$ is defined as the vector which maximizes (2.4), or equivalently, minimizes the negative log likelihood function

$$L(\underline{\beta}) = -\sum_{t=1}^{N} \log \xi_{k_t(1),\ldots,k_t(p_1),\underline{w}_{t,1},\ldots,\underline{w}_{t,p_2}}. \tag{2.5}$$

Under mild regularity conditions, it is well known that the ML estimate $\hat{\underline{\beta}}$ of $\underline{\beta}$ has the following nice statistical properties:

(1) $\hat{\underline{\beta}}$ is consistent and efficient;

(2) $\hat{\underline{\beta}}$ is asymptotically normal with mean $\underline{\beta}$ and covariance matrix equals to the inverse of the information matrix.

§2.3 Partition Maximum Likelihood Estimation

The classical maximum likelihood procedure is discussed in the previous section. Since the method requires a tremendous amount of computational time in evaluating the multivariate normal distribution probability as given in (2.2), direct minimization of $L(\underline{\beta})$ is not feasible, especially when p, the dimension of the random vector, is large. One method, similar to the Partition Maximum Likelihood (PML) method proposed by Poon and Lee (1987), to solve this problem is to compute the maximum likelihood estimates based on only a pair of variables, and then repeat with another pair until all the parameters of the whole model have been estimated. In this case, only single or double integrals are involved in the minimization, and hence save a lot of computational time. The basic idea of the procedure is described as follows.

We first partition the multivariate model into p(p-1)/2 bivariate sub-models. As two kinds of data are observed, we have to consider three cases,

(1) Polytomous-Polytomous pair,

(2) Polytomous-Interval pair,

(3) Interval-Interval pair.

Corresponding to different cases, we have different likelihood functions and sub-model parameters. We will deal with them respectively in §2.3.1 to §2.3.3. In §2.3.4, the asymptotic normality of the partition ML estimate $\bar{\underline{\beta}}$ of $\underline{\beta}$ is established.

§2.3.1 Polytomous-Polytomous Pair

Consider the bivariate sub-model corresponding to a pair of polytomous-polytomous pair (X_a, X_b), $a > b$, $a,b = 1,\ldots,p_1$, and let $\underline{\gamma}_{1,ab} = (\underline{\alpha}_a, \underline{\alpha}_b, \rho_{ab})'$ be the unknown parameter vector which consists of the thresholds, $\underline{\alpha}_i = (\alpha_{i,2},\ldots,\alpha_{i,n(i)})$, i= a,b; and the polychoric correlation ρ_{ab} of the bivariate model. The ML estimate $\hat{\underline{\gamma}}_{1,ab}$ of $\underline{\gamma}_{1,ab}$ is obtained by minimizing the following negative log-likelihood function,

$$
\begin{aligned}
L_{1,ab} &= L_{1,ab}(\underline{\gamma}_{1,ab}) \\
&= - \sum_{t=1}^{N} F_{1,ab,t},
\end{aligned}
$$

where
$$
\begin{aligned}
F_{1,ab,t} = \log\Big[&\Phi_2(\alpha_{t,a,k_t(a)+1}, \alpha_{t,b,k_t(b)+1}; \underline{0}, \underline{R}_{ab}) - \\
&\Phi_2(\alpha_{t,a,k_t(a)}, \alpha_{t,b,k_t(b)+1}; \underline{0}, \underline{R}_{ab}) - \\
&\Phi_2(\alpha_{t,a,k_t(a)+1}, \alpha_{t,b,k_t(b)}; \underline{0}, \underline{R}_{ab}) + \\
&\Phi_2(\alpha_{t,a,k_t(a)}, \alpha_{t,b,k_t(b)}; \underline{0}, \underline{R}_{ab}) \Big],
\end{aligned} \tag{2.6}
$$

with $\underline{R}_{ab} = \begin{bmatrix} 1.0 & \\ \rho_{ab} & 1.0 \end{bmatrix}$, is a function of $\underline{\alpha}_a$, $\underline{\alpha}_b$, and ρ_{ab}. Then, from the mean-value theorem, we have the following result.

<u>Lemma 2.1</u> If $\hat{\underline{\gamma}}_{1,ab}$ is the ML estimate of $\underline{\gamma}_{1,ab}$, then, under mild regularity conditions,

$$
N^{1/2}(\hat{\underline{\gamma}}_{1,ab} - \underline{\gamma}_{1,ab}) = - \underline{K}_{1,ab,ab}^{-1} N^{-1/2} \underline{\Delta}_{1,ab} + o_p(1) \tag{2.7}
$$

where $\underline{\Delta}_{1,ab} = \partial L_{1,ab}/\partial \underline{\gamma}_{1,ab}$ and $\underline{K}_{1,ab,ab}$ is the corresponding information matrix for $\underline{\gamma}_{1,ab}$.

<u>Proof</u>

Since $\hat{\underline{\gamma}}_{1,ab}$ is the ML estimate of $\underline{\gamma}_{1,ab}$, it must satisfy the following equation

$$
\frac{\partial \hat{L}_{1,ab}}{\partial \underline{\gamma}_{1,ab}} = \frac{\partial L_{1,ab}(\hat{\underline{\gamma}}_{1,ab})}{\partial \underline{\gamma}_{1,ab}} = \underline{0} . \tag{2.8}
$$

Then, it follows from the mean-value theorem, there exists a $\underline{\gamma}^*_{1,ab}$ between $\hat{\underline{\gamma}}_{1,ab}$ and $\underline{\gamma}_{1,ab}$ such that

$$N^{-1/2}\frac{\partial\hat{L}_{1,ab}}{\partial\underline{\gamma}_{1,ab}} = N^{-1/2}\frac{\partial L_{1,ab}}{\partial\underline{\gamma}_{1,ab}} + \left\{N^{-1}\frac{\partial^2 L_{1,ab}(\underline{\gamma}^*_{1,ab})}{\partial\underline{\gamma}_{1,ab}\partial\underline{\gamma}_{1,ab}}\right\}\left\{N^{1/2}(\hat{\underline{\gamma}}_{1,ab}-\underline{\gamma}_{1,ab})\right\}. \tag{2.9}$$

From (2.8), (2.9) may be rewritten as

$$N^{1/2}(\hat{\underline{\gamma}}_{1,ab} - \underline{\gamma}_{1,ab}) = - \left\{N^{-1}\frac{\partial^2 L_{1,ab}(\underline{\gamma}^*_{1,ab})}{\partial\underline{\gamma}_{1,ab}\partial\underline{\gamma}_{1,ab}}\right\}^{-1} \left\{N^{-1/2}\ \underline{\Delta}_{1,ab}\right\}.$$

As $\hat{\underline{\gamma}}_{1,ab}$ converges in probability to $\underline{\gamma}_{1,ab}$, we have $N^{-1}\frac{\partial^2 L_{1,ab}(\underline{\gamma}^*_{1,ab})}{\partial\underline{\gamma}_{1,ab}\partial\underline{\gamma}_{1,ab}}$ converges in probability to the information matrix $\underline{K}_{1,ab,ab}$. Therefore, we have

$$N^{1/2}(\hat{\underline{\gamma}}_{1,ab} - \underline{\gamma}_{1,ab}) = - \underline{K}^{-1}_{1,ab,ab}\ N^{-1/2}\ \underline{\Delta}_{1,ab} + o_p(1)$$

and complete the proof.■

§2.3.2 Polytomous-Interval Pair

Similar to the above case, we consider the bivariate sub-model (X_c, Y_d) with $c = 1,\ldots,p_1$, $d = 1,\ldots,p_2$ which is corresponding to a polytomous observation Z_c and an interval observation $\underline{w}_d=(w_{t,d,1}, w_{t,d,2})$. Let $\underline{\gamma}_{2,cd} = (\underline{\alpha}_c, \mu_d, \sigma_{dd}, \rho_{cd})'$ be the parameter vector of this sub-model. It can be noted that $\underline{\gamma}_{2,cd}$ consists of the thresholds of the polytomous variable Z_c, the mean and the variance of Y_d and the correlation of the two variables X_c and Y_d. The ML estimate $\hat{\underline{\gamma}}_{2,cd}$ of $\underline{\gamma}_{2,cd}$ is defined as the vector which minimizes the following negative log-likelihood function $L_{2,cd}(\underline{\gamma}_{2,cd})$,

$$L_{2,cd} = L_{2,cd}(\underline{\gamma}_{2,cd}) = -\sum_{t=1}^{N} F_{2,cd,t},$$

$$\text{where } F_{2,cd,t} = \log\Big[\Phi_2(\alpha_{t,c,k_t(c)+1}, w_{t,d,2}; \underline{\mu}_2, \underline{\Sigma}_{cd}) - \Phi_2(\alpha_{t,c,k_t(c)}, w_{t,d,2}; \underline{\mu}_2, \underline{\Sigma}_{cd}) - \Phi_2(\alpha_{t,c,k_t(c)+1}, w_{t,d,1}; \underline{\mu}_2, \underline{\Sigma}_{cd}) + \Phi_2(\alpha_{t,c,k_t(c)}, w_{t,d,1}; \underline{\mu}_2, \underline{\Sigma}_{cd})\Big], \tag{2.10}$$

with $\underline{\mu}_2 = (0, \mu_d)'$ and $\underline{\Sigma}_{cd} = \begin{bmatrix} 1.0 & \\ \rho_{cd} & \sigma_{dd} \end{bmatrix}$, is a function of $\underline{\alpha}_c$, μ_d, σ_{dd} and ρ_{cd}.

Follows from Lemma 2.1, it can be shown that

$$N^{1/2}(\hat{\underline{\gamma}}_{2,cd} - \underline{\gamma}_{2,cd}) = -\underline{K}^{-1}_{2,cd,cd}\, N^{-1/2}\, \underline{\Delta}_{2,cd} + o_p(1), \tag{2.11}$$

where $\underline{\Delta}_{2,cd} = \partial L_{2,cd} / \partial \underline{\gamma}_{2,cd}$ and $\underline{K}_{2,cd,cd}$ is the information matrix for $\underline{\gamma}_{2,cd}$.

§2.3.3 Interval-Interval Pair

Now, only the case of interval-interval variables is left behind. There are altogether $p_2(p_2-1)/2$ bivariate sub-models of this type. We consider the sub-model (Y_i,Y_j) which is corresponding to interval variables $\underline{W}_i$ and $\underline{W}_j$ with parameter vector $\underline{\gamma}_{3,ij} = \{\mu_i, \mu_j, \sigma_{ii}, \sigma_{jj}, \sigma_{ij}\}'$, $i,j=1,\ldots,p_2$, $i>j$. That is, the parameter vector $\underline{\gamma}_{3,ij}$ consists of the means, the variances and the covariance of Y_i and Y_j. Based on the data set D_t, the negative log-likelihood function can be expressed as follows:

$$L_{3,ij} = L_{3,ij}(\underline{\gamma}_{3,ij}) = - \sum_{t=1}^{N} F_{3,ij,t},$$

where

$$F_{3,ij,t} = \log\Big[\Phi_2(w_{t,i,2}, w_{t,j,2}; \underline{\mu}_3, \underline{\Sigma}_{ij}) - \Phi_2(w_{t,i,1}, w_{t,j,2}; \underline{\mu}_3, \underline{\Sigma}_{ij}) - \Phi_2(w_{t,i,2}, w_{t,j,1}; \underline{\mu}_3, \underline{\Sigma}_{ij}) + \Phi_2(w_{t,i,1}, w_{t,j,1}; \underline{\mu}_3, \underline{\Sigma}_{ij}) \Big], \tag{2.12}$$

with $\underline{\mu}_3 = (\ \mu_i,\ \mu_j)'$ and $\underline{\Sigma}_{ij} = \begin{bmatrix} \sigma_{ii} & \\ \sigma_{ij} & \sigma_{jj} \end{bmatrix}$, is a function of μ_i, μ_j, σ_{ii}, σ_{jj} and σ_{ij}. Then, similarly, from Lemma 2.1, the following result is established.

$$N^{1/2}(\hat{\underline{\gamma}}_{3,ij} - \underline{\gamma}_{3,ij}) = - \underline{K}_{3,ij,ij}^{-1} N^{-1/2} \underline{\Delta}_{3,ij} + o_p(1), \tag{2.13}$$

where $\underline{\Delta}_{3,ij} = \partial L_{3,ij} / \partial \underline{\gamma}_{3,ij}$ and $\underline{K}_{3,ij,ij}$ is the information matrix for $\underline{\gamma}_{3,ij}$.

§2.3.4 Normality of the Partition Estimators

There are totally $p_1(p_1-1)/2$ polytomous-polytomous sub-models, p_1p_2 polytomous-interval sub-models and $p_2(p_2-1)/2$ interval-interval sub-models. Each parameter vector of these sub-models satisfies (2.7), (2.11) and (2.13) respectively. By letting

$$\underline{\gamma} = (\underline{\gamma}'_{1,21}, \ldots, \underline{\gamma}'_{1,p_1(p_1-1)}, \underline{\gamma}'_{2,11}, \ldots, \underline{\gamma}'_{2,p_1p_2}, \underline{\gamma}'_{3,21}, \ldots, \underline{\gamma}'_{3,p_2(p_2-1)})'$$

and

$$\underline{\Delta}_{\gamma} = \frac{\partial L}{\partial \underline{\gamma}} = \left[\frac{\partial L'_{1,21}}{\partial \underline{\gamma}_{1,21}}, \ldots, \frac{\partial L'_{1,p_1(p_1-1)}}{\partial \underline{\gamma}_{1,p_1(p_1-1)}}, \frac{\partial L'_{2,11}}{\partial \underline{\gamma}_{2,11}}, \ldots, \frac{\partial L'_{2,p_1p_2}}{\partial \underline{\gamma}_{2,p_1p_2}}, \frac{\partial L'_{3,21}}{\partial \underline{\gamma}_{3,21}}, \ldots, \frac{\partial L_{3,p_2(p_2-1)}}{\partial \underline{\gamma}_{3,p_2(p_2-1)}} \right]',$$

we can integrate the equations of the form (2.7), (2.11) and (2.13) into a matrix, and get the following result,

$$N^{1/2} (\hat{\underline{\gamma}} - \underline{\gamma}) = - \underline{K}^{-1} (N^{-1/2} \underline{\Delta}_{\gamma}) + o_p(1) \tag{2.14}$$

where $\underline{K}$ is a diagonal block matrix of the form,

$$\underline{K} = \begin{bmatrix} \underline{K}_{1,21,21} & & & & & & & \\ & \ddots & & & & & \underline{0} & \\ & & \underline{K}_{1,p_1(p_1-1),p_1(p_1-1)} & & & & & \\ & & & \underline{K}_{2,11,11} & & & & \\ & & & & \ddots & & & \\ & & & & & \underline{K}_{2,p_1p_2,p_1p_2} & & \\ & \underline{0} & & & & & \underline{K}_{3,21,21} & \\ & & & & & & & \ddots \\ & & & & & & & \underline{K}_{3,p_2(p_2-1),p_2(p_2-1)} \end{bmatrix}.$$

<u>Theorem 2.1</u> The asymptotic distribution of $N^{1/2} (\hat{\underline{\gamma}} - \underline{\gamma})$ is multivariate normal with zero mean and covariance matrix $\underline{K}^{-1}\underline{\nabla}_{\gamma}\underline{K}^{-1}$ where $\underline{\nabla}_{\gamma}$ is the asymptotic covariance matrix of $N^{-1/2}\ \partial L / \partial \underline{\gamma}$.

<u>Proof</u>

Since $\underline{\Delta}_{1,ab}$, for $a,b=1,\ldots,p_1$, $a>b$; $\underline{\Delta}_{2,ab}$, for $a=1,\ldots,p_1$, $b=1,\ldots,p_2$; and $\underline{\Delta}_{3,ab}$ for $a,b=1,\ldots,p_2$, $a>b$ are functions of independent and identical distributed observations from the data set D_t and by property of ML estimates, $E(\underline{\Delta}_{k,ab}) = \underline{0}$ for $k = 1, 2, 3$, and hence, $E(\underline{\Delta}_{\gamma}) = \underline{0}$, we will obtain, by central limit theorem,

$$N^{-1/2}\frac{\partial L}{\partial \underline{\gamma}} \overset{L}{===} N[\underline{0}, \underline{\nabla}_{\underline{\gamma}}]$$

where " $\overset{L}{===}$ " means convergent in distribution and $\underline{\nabla}_{\underline{\gamma}}$ is the asymptotic covariance matrix. It follows from (2.14), the following result is established.

$$N^{1/2} (\hat{\underline{\gamma}} - \underline{\gamma}) \overset{L}{===} N[\underline{0}, \underline{K}^{-1}\underline{\nabla}_{\underline{\gamma}} \underline{K}^{-1}]. \blacksquare \tag{2.15}$$

It can be noticed that the parameters $\underline{\alpha}$, $\underline{\mu}_y$, and $\underline{\sigma}_{ii}$ with $i = p_1+1,\dots,p$, contained in $\underline{\beta}$ have been estimated for (p - 1) times and ρ_{ij}, $i,j=1,\dots,p_1$, $i>j$; σ_{ij}, $i = 1,\dots,p_1$, $j = p_1+1,\dots,p$; σ_{ij}, $i, j = p_1+1,\dots,p$, $i>j$; for 1 time . Let $\bar{\underline{\beta}}$ be the partitioned maximum likelihood (PML) estimate of $\underline{\beta}$ which is obtained by averaging the corresponding ML estimates of $\hat{\underline{\gamma}}$ from the bivariate sub-models. Then there exists an appropriate selection matrix $\underline{S}_{\underline{\beta}}$ which picks up the correct weighted components of the $\hat{\underline{\gamma}}$'s from the appropriate bivariate sub-models to form the PML parameter estimates, $\bar{\underline{\beta}}$. By applying multivariate normal distribution theory to (2.15), we can obtain the asymptotic distribution of the consistent estimate, $\bar{\underline{\beta}}$, of $\underline{\beta}$ which is given by

$$N^{1/2} (\bar{\underline{\beta}} - \underline{\beta}) \overset{L}{===} N[\underline{0}, \underline{\Omega}_{\bar{\underline{\beta}}}],$$

where $\underline{\Omega}_{\bar{\underline{\beta}}} = \underline{S}_{\underline{\beta}} \underline{K}^{-1} \underline{\nabla}_{\underline{\gamma}} \underline{K}^{-1} \underline{S}'_{\underline{\beta}}$. Thus, the asymptotic normality of $\bar{\underline{\beta}}$ is derived.

§2.4 Optimization Procedure and Simulation Study

§2.4.1 Optimization Procedure

As mentioned above, the maximum likelihood estimate of the parameter vector for each bivariate sub-model is obtained by minimizing its corresponding negative log-likelihood function. In practice, the minimum of the negative log-likelihood function cannot be obtained in closed form. Hence, some iterative algorithms such as Newton-Raphson or Fisher-Scoring Algorithm (See, e.g., Lee & Jennrich,1979) should be used.

The procedure for minimizing (2.6) for polytomous-polytomous pair has been developed by Poon & Lee (1987) and a program based on Fisher-Scoring algorithm written in FORTRAN IV with double precision has been implemented. For each (a,b) pair, a, b $=1,\ldots,p_1$, a>b; the basic step of Scoring algorithm is given by

$$\Delta\underline{\gamma}_1 = - \zeta\, \underline{K}_1(\underline{\gamma}_1)^{-1}\, \dot{L}_1(\underline{\gamma}_1)$$

where ζ is a step-size parameter which takes the first value in the sequence { 1, 1/2, 1/4,...} that reduces $L_1(\underline{\gamma}_1)$,

$$\dot{L}_1(\underline{\gamma}_1) = \partial L_1(\underline{\gamma}_1) \,/\, \partial\underline{\gamma}_1$$

$$= - \sum_{t=1}^{N} \frac{1}{\xi_{1,t}} \frac{\partial\xi_{1,t}}{\partial\underline{\gamma}_1}$$

is the gradient vector and $\underline{K}_1(\underline{\gamma}_1)$ is the information matrix

$$\underline{K}_1(\underline{\gamma}_1) = E\left(\frac{\partial L_1(\underline{\gamma}_1)}{\partial \underline{\gamma}_1} \frac{\partial L_1(\underline{\gamma}_1)}{\partial \underline{\gamma}_1}'\right)$$

$$= N \sum_{t=1}^{N} \frac{1}{\xi_{1,t}} \frac{\partial \xi_{1,t}}{\partial \underline{\gamma}_1} \frac{\partial \xi_{1,t}}{\partial \underline{\gamma}_1}' ,$$

where $\xi_{1,t}$ is the probability that a particular observation of the polytomous variables Z_a and Z_b for subject t is taken. That is,

$$\xi_{1,t} = \Big[\Phi_2(\alpha_{t,a,k_t(a)+1}, \alpha_{t,b,k_t(b)+1}; \underline{0}, \underline{R}_{ab}) - \Phi_2(\alpha_{t,a,k_t(a)}, \alpha_{t,b,k_t(b)+1}; \underline{0}, \underline{R}_{ab}) - \Phi_2(\alpha_{t,a,k_t(a)+1}, \alpha_{t,b,k_t(b)}; \underline{0}, \underline{R}_{ab}) + \Phi_2(\alpha_{t,a,k_t(a)}, \alpha_{t,b,k_t(b)}; \underline{0}, \underline{R}_{ab})\Big].$$

And we can use $\|\dot{L}_1(\underline{\gamma}_1)\|$, the root mean square of elements in $\dot{L}_1(\underline{\gamma}_1)$, as the convergence criterion and we stop if the root mean squares is smaller than a pre-assigned value ϵ. Hence, we need the explicit expression of $\partial\xi_{1,t}/\partial\underline{\gamma}_1$ to implement the algorithm. As $\Phi_2(\alpha_1,\alpha_2;\underline{0},\underline{\Sigma})$ is involved in $\xi_{1,t}$, the gradient vector and the information matrix involves partial derivations of the form $\partial\Phi_2(\alpha_1,\alpha_2;\underline{0},\underline{\Sigma})/\partial\alpha_i$, and $\partial\Phi_2(\alpha_1,\alpha_2;\underline{0},\underline{\Sigma})/\partial\rho_{ij}$. The details of the derivatives were derived by Poon and Lee (1987) and Johnson and Kotz (1972). Since, $\underline{K}_1(\hat{\underline{\gamma}}_1)$, the information matrix evaluated at the final Scoring estimate, converages in probability to the information matrix, we can use $\underline{K}_1(\hat{\underline{\gamma}}_1)$ to approximate the information matrix. As a result, the scoring algorithm produces not only the maximum likelihood estimate but also an approximation of its asymptotic covariance matrix and hence its stardard errors.

In practice, the Scoring algorithm is found very unstable when interval data is used if the sample size N is too small, say, smaller than 400. Hence, another more efficient but more complicated in implementation, Newton-Raphson algorithm, is used for the rest two cases. Its basic step is summarized as follows:

$$\Delta\underline{\gamma}_k = -\zeta\, \underline{H}_k(\underline{\gamma}_k)^{-1}\, \dot{L}_k(\underline{\gamma}_k) \text{ for } k = 2,3$$

where ζ and $\dot{L}_k(\underline{\gamma}_k)$ are the step-size parameter and the gradient vector as in the Scoring algorithm whereas

$$\underline{H}_k(\underline{\gamma}_k) = \frac{\partial^2 L_k(\underline{\gamma}_k)}{\partial\underline{\gamma}_k \partial\underline{\gamma}_k}$$

is the Hessian matrix. In these cases, the same stopping criterion is used. Therefore, to implement the Newton-Raphson algorithm, we need the explicit expression of $\dot{L}_k(\underline{\gamma}_k)$ and $\underline{H}_k(\underline{\gamma}_k)$. In general, it is necessary to find $\partial\Phi_2(\alpha, \beta;\ \underline{\mu},\ \underline{\Sigma}) \ /\ \partial\underline{\gamma}_k$ and $\partial^2\Phi_2(\alpha,\beta;\ \underline{\mu},\ \underline{\Sigma}) \ /\ \partial\underline{\gamma}_k\partial\underline{\gamma}_k$. The expressions for these derivations can be found in Poon & Lee(1987) and Johnson and Kotz(1972) as before. Since the Hessian matrix $\underline{H}_k$ converges in probability to its corresponding information matrix $\underline{K}_k$, we can use the Hessian matrix $\underline{H}_k$ to approximate the information matrix $\underline{K}_k$. Therefore, similar to the Scoring algorithm presented before, the Newton-Raphson algorithm also produces the maximum likelihood estimate and an approximation of its asymptotic covariance matrix and hence standard errors.

§2.4.2 Simulation Study

Based on the algorithm discussed in the previous section, a computer program written in FORTRAN IV with double precision has been implemented to obtain the PML estimates. To study the behavior of the estimator under different situations, monte carlo studies have been conducted. The performance of the estimates under different sample sizes, diffenent types of intervals for interval variables and different sets of threshold values for polytomous variables are discussed.

The study is based on simulated data from a multivariate normal distribution with the dimensions of $\underline{X}$ and $\underline{Y}$ are both two. The population mean vector $\underline{\mu}_y = (\mu_1, \mu_2)'$ of $\underline{Y} = (Y_1, Y_2)'$ is taken to be $\underline{0}$ and the population covariance matrix $\underline{\Sigma}$ and the threshold values are taken as follows:

$$\underline{\Sigma} = \begin{bmatrix} 1.0 & & & \\ \rho_{21} & 1.0 & & \\ \rho_{31} & \rho_{32} & \sigma_{33} & \\ \rho_{41} & \rho_{42} & \sigma_{43} & \sigma_{44} \end{bmatrix} = \begin{bmatrix} 1.0 & & & \\ 0.5 & 1.0 & & \\ 0.5 & 0.5 & 1.0 & \\ 0.5 & 0.5 & 0.5 & 1.0 \end{bmatrix};$$

and

(a) Symmetric distribution of Z_1 and Z_2:

$$\underline{\alpha}_1 = \{ \alpha_{1,1} = -\infty, \alpha_{1,2} = -0.5, \alpha_{1,3} = 0.5, \alpha_{1,4} = \infty \},$$

$$\underline{\alpha}_2 = \{ \alpha_{2,1} = -\infty, \alpha_{2,2} = -0.5, \alpha_{2,3} = 0.5, \alpha_{2,4} = \infty \},$$

(b) Asymmetric distribution of Z_1 and Z_2:

$$\underline{\alpha}_1 = \{ \alpha_{1,1} = -\infty, \alpha_{1,2} = 0.0, \alpha_{1,3} = 1.0, \alpha_{1,4} = \infty \},$$

$$\underline{\alpha}_2 = \{ \alpha_{2,1} = -\infty, \alpha_{2,2} = 0.0, \alpha_{2,3} = 0.5, \alpha_{2,4} = \infty \}.$$

Multivariate normal variates were generated by the IMSL(1975) subroutine RNMVN with the specified mean vector and covariance matrix $\underline{\Sigma}$. The interval observations were created by addition and subtraction of small numbers and the polytomous observations were created via (2.1). For example, we might subtract 0.3 from and add 0.1 to a generated normal variate 0.25 to obtain an interval observation (-0.05, 0.35) or classified 0.25 as second category if the threshold values $\underline{\alpha}_1$ in (a) is used. Thus, in order to create interval variates, the following criterion is used.

(i) Small length of the interval

proportion of variates	add	subtract
1/4	0.03	0.01
1/2	0.04	0.02
1/4	0.01	0.03

(ii) Intermediate length of the interval

proportion of variates	add	subtract
1/4	0.3	0.1
1/2	0.2	0.2
1/4	0.1	0.3

We have studied all 4 combinations of the different threshold values and length of the interval for interval observations with different sample sizes N = 70, 100, 200, 300, and 400. For each combination, 40 replications were generated and the convergence criterion ϵ is taken to be 0.00005. The simulation results concerning the mean vector, the polychoric correlation, the correlation between polytomous and interval variables, the variances for the interval variables as well as the threshold estimates are presented in

Tables 1. They are denoted by 1a(i), 1b(i), 1a(ii) and 1b(ii). (Note that 1a(i) refers to the case that the threshold values in (a) and the interval length in (i) are used). In the tables, the following statistics are given.

(1) The mean of the estimates

$$\bar{\bar{\beta}}_i = \frac{1}{40}\left\{ \sum_{k=1}^{40} \bar{\beta}_i(k) \right\},$$

where $\bar{\beta}_i(k)$ is the i-th element of $\bar{\underline{\beta}}$ in the k-th replication.

(2) The sample standard errors of the estimates

$$\overline{S.E.}_i = \left\{ \frac{1}{39} \sum_{k=1}^{40} (\bar{\beta}_i(k) - \bar{\bar{\beta}}_i)^2 \right\}^{1/2}.$$

(3) The average of estimated standard errors of the estimates

$$S.E._i = \frac{1}{40}\left\{ \sum_{k=1}^{40} \text{Standard error of } (\bar{\beta}_i(k)) \right\}.$$

(4) The ratio of the sample standard errors to the average of estimated standard errors of the estimates

$$R_i = \overline{S.E.}_i \, / \, S.E._i.$$

We would expect that $\overline{S.E.}_i$ is close to $S.E._i$ and thus the ratio R_i would be close to 1.

(5) The root-mean-squared errors

$$\mathrm{RMS} = \left\{ \frac{1}{40} \sum_{k=1}^{40} (\bar{\beta}_i(k) - \beta_i)^2 \right\}^{1/2}.$$

where β_i is the true value of the i-th element of the parameter vector.

From the Tables 1, the followings are observed:

(1) The means of the estimates are very close to the true values and the RMSs are very small in all situations.

(2) In many cases, the sample size increases, then the RMS decreases. However, the RMSs are not very stable in the case of small samples. For example, the increase in RMSs always occurs in the case of sample sizes 70 and 100, which contributes 13 out of 31 unsatisfactory observations.

(3) The Ratios fall into the range (0.8, 1.2) in all cases except for the correlation of the two interval variables. This indicates that estimates for the standard errors are acceptable. The large fluctuations of Ratio's for the correlation of the two interval variables may be due to the fuzziness of the data that the thresholds may be different for different subject.

(4) The method tends to underestimate the standard errors for the estimates for the correlation of the two interval variables but the situation becomes better as the sample size N increases.

Chapter 3

Three-stage Procedure for Covariance Structure Analysis

In the previous chapter, the partition maximum likelihood estimate of the correlation between polytomous and interval data is studied. Now, the structural equation models with polytomous and interval data will be discussed. The main purpose of this chapter is to develop basic theories for analysis of covariance structure with these two kinds of data. A three-stage procedure is employed. In the first stage, the estimates of the thresholds for the polytomous variables are obtained via the partition maximum likelihood method. In the second stage, the pseudo maximum likelihood method is used to obtain the estimates of the elements in the covariance matrix, $\underline{\Sigma}$, without imposing any structure and the thresholds are fixed at the estimates in stage one. It will be shown that the joint asymptotic distribution of the estimates of the elements in $\underline{\Sigma}$ is multivariate normal. The third stage estimates the structure via the generalized least squares approach with an appropriate weight matrix given in stage two. Asymptotic properties for further statistical inference of the model, such as the asymptotic distribution of the structural parameter estimates and the goodness-of-fit statistic, will be developed. A Monte Carlo study will be conducted to study the performance of the estimates.

§ 3.1 Model

We consider the general multivariate covariance model with polytomous and interval variables as given in Chapter 2. The joint distribution of the two latent continuous vectors $\underline{X}$ and $\underline{Y}$ with dimensions p_1 and p_2 is assumed to be multivariate normal with mean vector $(\underline{0}', \underline{\mu}_y')'$ and covariance matrix

$$\underline{\Sigma} = \underline{\Sigma}(\underline{\theta}_o) = \begin{bmatrix} \underline{\Sigma}_{xx} & \underline{\Sigma}_{xy} \\ \underline{\Sigma}_{yx} & \underline{\Sigma}_{yy} \end{bmatrix},$$

where $\underline{\theta}_o$ is a $q \times 1$ structural parameter vector. It is assumed that $\underline{\Sigma} = \underline{\Sigma}(\underline{\theta})$ is identified, that is, $\underline{\Sigma}(\underline{\theta}_1) = \underline{\Sigma}(\underline{\theta}_2)$ implies $\underline{\theta}_1 = \underline{\theta}_2$.

Suppose now that $\underline{X}$ can only be observed through a polytomous random vector $\underline{Z} = (Z_1, \ldots, Z_{p_1})$ and their relationship is given by

$$Z_i = k(i) \quad \text{if } \alpha_{i,k(i)} \le X_i < \alpha_{i,k(i)+1},$$

for $k(i) = 1, \ldots, n(i)$, $i = 1, \ldots, p_1$; where $\{\alpha_{i,1} = -\infty, \alpha_{i,2}, \ldots, \alpha_{i,n(i)}, \alpha_{i,n(i)+1} = \infty\}$ are the thresholds and $n(i)$ is the number of categories corresponding to the i-th variable of $\underline{Z}$. For each random variable Y_j, $j = 1, \ldots, p_2$, the corresponding observation is an interval $(w_{j,1}, w_{j,2})$ rather than a point with $w_{j,1} < w_{j,2}$.

§ 3.2 Three-stage Estimation Method

Suppose we observe a random sample of polytomous and interval data with sample size N which is given by

$$D_t = \left[k_t(1), \ldots, k_t(p_1), \underline{w}_{t,1}, \ldots, \underline{w}_{t,p_2} \right],$$

where $t = 1, \ldots, N$, is the subject index. Similar in Chapter 2, we also define $\alpha_{t,i,k_t(i)} = \alpha_{i,k(i)}$ if $k_t(i) = k(i)$ for $i = 1,\ldots, p_1$.

In Chapter 2, it is shown that the parameters, thresholds, means, covariances and variances, of the polytomous variables cannot be estimated simultaneously. Thus, in general, to handle this identification problem, the diagonal elements of $\underline{\Sigma}_{xx}$, diag($\underline{\Sigma}_{xx}$), and the mean vector $\underline{\mu}_x$ will be fixed to the identity matrix and zero vector respectively. In this case, only the correlations can be estimated, and hence, only the correlation structure can be considered which will be discussed in next chapter. However, if the thresholds are known, the covariance and the variances of the polytomous variables can be identified. Then, the covariance structure model can be considered. Hence, it is reasonable to consider the pseudo maximum likelihood approach (See, Parke (1986), Lee, Poon, and Bentler (1990b)) to handle the covariance structure model. The classical maximum likelihood method will not be discussed here since the procedure is time consuming as mentioned in the previous chapter. Instead, the partition maximum likelihood method is considered.

The basic idea of the three-stage procedure is first to partition the multivariate model into p(p-1)/2 sub-models. In the first stage, only the polytomous-polytomous pairs are considered and all the thresholds are estimated in this stage. Then, the estimates of the thresholds obtained in first stage are treated as given. All the p(p-1)/2 bivariate sub-models are used to estimate the elements in the covariance matrix $\underline{\Sigma}$ without imposing any structure in the second stage. Moreover, an appropriate weight matrix will also be given. The final stage estimates the structural parameter vector $\underline{\theta}_o$ by using generalized least squares method. The procedure is briefly described as follows:

Stage I :

Follow the method as in Poon and Lee (1987) or the procedure discussed in Chapter 2, we can obtain the consistent estimates, $\bar{\underline{\alpha}}$, of the thresholds $\underline{\alpha}$ and it can be shown that $\bar{\underline{\alpha}}$ is jointly asymptotically multivariate normally distributed, that is,

$$N^{1/2} (\bar{\underline{\alpha}} - \underline{\alpha}) \overset{L}{===} N[\underline{0}, \Omega_{\bar{\underline{\alpha}}}], \qquad (3.1)$$

where $\Omega_{\bar{\underline{\alpha}}}$ is the corresponding asymptotic covariance matrix.

Stage II :

The thresholds of the model at this stage will be fixed at $\bar{\underline{\alpha}}$. It should be noted that by fixing the thresholds, the dispersions for the polytomous variables are identified. Therefore, we can estimate the variances of polytomous variables, and hence, the means, variances, and covariances are taken as the parameters in this stage and denoted by $\underline{\psi}$ = ({ μ_y}, $\underline{\sigma}$ = { σ_{ij}, i,j = 1,...,p and i≥j }). These parameters are estimated by the pseudo maximum likelihood approach, giving a consistent estimate $\bar{\underline{\Sigma}}$ of $\underline{\Sigma}$.

It is well known that computational algorithms would be more stable with less parameters. By fixing the thresholds, number of parameters involved in different submodels are reduced. If the number of categories for the polytomous variables are large, this reduction is significant and would improve substantially the stability of the algorithms, especially for the polytomous-interval pair. With these advantages, less computer time is required to obtain the estimates.

The normality of the pseudo maximum likelihood estimates will be derived in the following. Similar to Chapter 2, we have to consider 3 different cases.

(1) Polytomous-Polytomous Pair

Consider the bivariate sub-model (X_a, X_b), a, b = 1,..., p_1, a>b with corresponding polytomous variables Z_a and Z_b. Let $\underline{\alpha}_{ab} = (\underline{\alpha}_a', \underline{\alpha}_b')' = (\{\alpha_{a,2}, \ldots, \alpha_{a,n(a)}\}, \{ \alpha_{b,2}, \ldots, \alpha_{b,n(b)} \})'$ and $\underline{\beta}_{1,ab} = \{ \sigma_{aa}, \sigma_{bb}, \sigma_{ab} \}'$. Suppose $\bar{\underline{\alpha}}_{ab}$ is a consistent estimate of $\underline{\alpha}_{ab}$, for example, $\bar{\underline{\alpha}}_{ab}$ can be obtained in stage one discussed before. Then, the pseudo maximum likelihood estimates $\hat{\underline{\beta}}_{1,ab}(\bar{\underline{\alpha}}_{ab})$ of $\underline{\beta}_{1,ab}$ are defined as the vector which maximizes the corresponding likelihood function, or equivalently, minimizes the negative log-likelihood function,

$$\begin{aligned} L_{1,ab} &= L_{1,ab}(\bar{\underline{\alpha}}_{ab}, \underline{\beta}_{1,ab}) \\ &= -\sum_{t=1}^{N} F_{1,ab,t}, \end{aligned}$$

$$\begin{aligned} \text{where } F_{1,ab,t} = \log \Big[& \Phi_2(\bar{\alpha}_{t,a,k_t(a)+1}, \bar{\alpha}_{t,b,k_t(b)+1}; \underline{0}, \underline{\Sigma}_{ab}) - \\ & \Phi_2(\bar{\alpha}_{t,a,k_t(a)}, \bar{\alpha}_{t,b,k_t(b)+1}; \underline{0}, \underline{\Sigma}_{ab}) - \\ & \Phi_2(\bar{\alpha}_{t,a,k_t(a)+1}, \bar{\alpha}_{t,b,k_t(b)}; \underline{0}, \underline{\Sigma}_{ab}) + \\ & \Phi_2(\bar{\alpha}_{t,a,k_t(a)}, \bar{\alpha}_{t,b,k_t(b)}; \underline{0}, \underline{\Sigma}_{ab}) \Big], \end{aligned} \tag{3.2}$$

with $\underline{\Sigma}_{ab} = \begin{pmatrix} \sigma_{aa} & \\ \sigma_{ab} & \sigma_{bb} \end{pmatrix}$ and t is the subject index. Under the conditions given in Gong & Samaniego (1981), the following lemmas are obtained.

Lemma 3.1 Under mild regularity conditions, the second derivatives

$$N^{-1} \frac{\partial^2 L_{1,ab}(\bar{\underline{\alpha}}_{ab}, \underline{\beta}_{1,ab})}{\partial \underline{\beta}_{1,ab} \; \partial \underline{\beta}_{1,ab}} = \underline{I}_{1,ab,ab} + o_p(1),$$

where $\underline{I}_{1,ab,ab}$ is the information matrix for $\underline{\beta}_{1,ab}$ evaluated at the true values of $\underline{\alpha}_{ab}$ and $\underline{\beta}_{1,ab}$.

Proof : See Gong and Samaniego, 1981.■

Lemma 3.2 The pseudo ML estimate $\hat{\underline{\beta}}_{1,ab}(\bar{\underline{\alpha}}_{ab})$ satisfies the following system of equations

$$N^{1/2} (\hat{\underline{\beta}}_{1,ab}(\bar{\underline{\alpha}}_{ab}) - \underline{\beta}_{1,ab}) = - \underline{I}_{1,ab,ab}^{-1} N^{-1/2} \bar{\underline{\Delta}}_{1,ab} + o_p(1), \tag{3.3}$$

where $\bar{\underline{\Delta}}_{1,ab}(\bar{\underline{\alpha}}_{ab}, \underline{\beta}_{1,ab}) = \partial L_{1,ab}/\partial \underline{\beta}_{1,ab}$ evaluated at $\bar{\underline{\alpha}}_{ab}$ and $\underline{I}_{1,ab,ab}$ is the corresponding information matrix for $\underline{\beta}_{1,ab}$.

Proof : By Lemma 2.1 expanding on $\underline{\beta}_{1,ab}$ and Lemma 3.1, (3.3) consequently holds.■

(2)Polytomous-Interval Pair

The bivariate sub-model corresponding to (X_a, Y_b) for $a = 1,\dots,p_1$; $b = 1,\dots,p_2$ is considered. As the interval variable $\underline{W}_b$ is involved in this sub-model, the mean parameter $\mu_{y,b}$ should be involved. Hence, the parameter vector for the sub-model becomes $\underline{\beta}_{2,ab} = (\mu_{y,b}, \sigma_{aa}, \sigma_{bb}, \sigma_{ab})'$ and the thresholds, $\underline{\alpha}_a$, of the polytomous variable, Z_a, becomes the nuisance parameter vector. Let $\underline{\mu}_2 = (0, \mu_{y,b})'$ and $\underline{\Sigma}_{ab}$ be the covariance matrix with σ_{aa}, σ_{bb}, σ_{ab} as its elements. Then, based on the data set D_t, the negative log-likelihood function is defined as

$$
\begin{aligned}
L_{2,ab} &= L_{2,ab}(\bar{\underline{\alpha}}_a, \underline{\beta}_{2,ab}) \\
&= - \sum_{t=1}^{N} F_{2,ab,t},
\end{aligned}
$$

$$
\begin{aligned}
\text{where } F_{2,ab,t} = \log\Big[&\Phi_2(\bar{\alpha}_{t,a,k_t(a)+1}, w_{t,b,2}; \underline{\mu}_2, \underline{\Sigma}_{ab}) - \\
&\Phi_2(\bar{\alpha}_{t,a,k_t(a)}, w_{t,b,2}; \underline{\mu}_2, \underline{\Sigma}_{ab}) - \\
&\Phi_2(\bar{\alpha}_{t,a,k_t(a)+1}, w_{t,b,1}; \underline{\mu}_2, \underline{\Sigma}_{ab}) + \\
&\Phi_2(\bar{\alpha}_{t,a,k_t(a)}, w_{t,b,1}; \underline{\mu}_2, \underline{\Sigma}_{ab})\Big],
\end{aligned} \tag{3.4}
$$

and t is the subject index.

Let $\hat{\underline{\beta}}_{2,ab}(\bar{\underline{\alpha}}_a)$ be the pseudo ML estimates for $\underline{\beta}_{2,ab}$, then by following the similar arguments for the polytomous-polytomous pair, the following result is obtained.

$$
N^{1/2} (\hat{\underline{\beta}}_{2,ab}(\bar{\underline{\alpha}}_a) - \underline{\beta}_{2,ab}) = - \underline{I}_{2,ab,ab}^{-1} N^{-1/2} \bar{\underline{\Delta}}_{2,ab} + o_p(1), \tag{3.5}
$$

where $\bar{\underline{\Delta}}_{2,ab}(\bar{\underline{\alpha}}_a, \underline{\beta}_{2,ab}) = \partial L_{2,ab}/\partial \underline{\beta}_{2,ab}$ evaluated at $\bar{\underline{\alpha}}_a$ and $\underline{I}_{2,ab,ab}$ is the corresponding information matrix.

(3) Interval-Interval Pair

Consider the bivariate sub-model corresponding to a pair of interval-interval pair (Y_a, Y_b), a, b = 1, ..., p_2; a>b. It should be noticed that no polytomous variable is involved in the sub-model, thus, no pseudo estimates are involved in this case. Let $\underline{\beta}_{3,ab} = (\mu_{y,a}, \mu_{y,b}, \sigma_{aa}, \sigma_{bb}, \sigma_{ab})'$ be the parameter vector, then the negative log-likelihood function is defined as

$$L_{3,ab} = L_{3,ab}(\underline{\beta}_{3,ab}) = -\sum_{t=1}^{N} F_{3,ab,t},$$

where $F_{3,ab,t}=\log\Big[\Phi_2(w_{t,a,2}, w_{t,b,2}; \underline{\mu}_3, \underline{\Sigma}_{ab}) - \Phi_2(w_{t,a,1}, w_{t,b,2}; \underline{\mu}_3, \underline{\Sigma}_{ab}) - \Phi_2(w_{t,a,2}, w_{t,b,1}; \underline{\mu}_3, \underline{\Sigma}_{ab}) + \Phi_2(w_{t,a,1}, w_{t,b,1}; \underline{\mu}_3, \underline{\Sigma}_{ab})\Big]$, (3.6)

is a function of $\underline{\mu}_3 = (\mu_{y,a}, \mu_{y,b})'$ and $\underline{\Sigma}_{ab} = \begin{bmatrix} \sigma_{aa} & \\ \sigma_{ab} & \sigma_{bb} \end{bmatrix}$, and t is the subject index.

Then, follows the arguments used in interval-interval pair in Chapter 2, an analogous result to (2.13) is true for the ML estimates $\hat{\underline{\beta}}_{3,ab}$ of $\underline{\beta}_{3,ab}$ and namely,

$$N^{1/2}(\hat{\underline{\beta}}_{3,ab} - \underline{\beta}_{3,ab}) = -\underline{I}_{3,ab,ab}^{-1} N^{-1/2} \underline{\Delta}_{3,ab} + o_p(1), \tag{3.7}$$

where $\underline{\Delta}_{3,ab}(\underline{\beta}_{3,ab}) = \partial L_{3,ab}/\partial \underline{\beta}_{3,ab}$ and $\underline{I}_{3,ab,ab}$ is the corresponding information matrix.

As the estimates of the parameter vectors can be expressed in similar form in this stage, (3.3), (3.5) and (3.7) can be rewritten into matrix form by denoting

$$\underline{\beta} = (\underline{\beta}_{1,21}', \ldots, \underline{\beta}_{1,p_1(p_1-1)}', \underline{\beta}_{2,11}', \ldots, \underline{\beta}_{2,p_1p_2}', \underline{\beta}_{3,21}', \ldots, \underline{\beta}_{3,p_2(p_2-1)}')'$$

and

$$\underline{\Delta}_{\underline{\beta}} = \frac{\partial L}{\partial \underline{\beta}} = \left[\frac{\partial L_{1,21}'}{\partial \underline{\beta}_{1,21}}, \ldots, \frac{\partial L_{1,p_1(p_1-1)}'}{\partial \underline{\beta}_{1,p_1(p_1-1)}}, \frac{\partial L_{2,11}'}{\partial \underline{\beta}_{2,11}}, \ldots, \frac{\partial L_{2,p_1p_2}'}{\partial \underline{\beta}_{2,p_1p_2}}, \frac{\partial L_{3,21}'}{\partial \underline{\beta}_{3,21}}, \ldots, \frac{\partial L_{3,p_2(p_2-1)}'}{\partial \underline{\beta}_{3,p_2(p_2-1)}} \right]', \tag{3.8}$$

we have,

$$N^{1/2} (\hat{\underline{\beta}}(\bar{\underline{\alpha}}) - \underline{\beta}) = - \underline{I}^{-1} (N^{-1/2} \underline{\Delta}_{\underline{\beta}}(\bar{\underline{\alpha}})) + o_p(1), \tag{3.9}$$

where $\underline{I}$ is a diagonal block matrix with diagonal blocks $\underline{I}_{1,ij,ij}$ for $i,j=1,\ldots,p_1$, $i>j$; $\underline{I}_{2,ij,ij}$ for $i = 1,\ldots,p_1$, $j = 1,\ldots,p_2$; and $\underline{I}_{3,ij,ij}$ for $i,j = 1,\ldots,p_2$, $i>j$. That is, $\underline{I}$ will take the form as $\underline{K}$ in Chapter 2.

It can be noticed that $\dfrac{\partial L_{k,ij}(\bar{\underline{\alpha}})}{\partial \underline{\beta}_{k,ij}}$ depend on the estimates of the thresholds, $\bar{\underline{\alpha}}$, obtained in stage one when $k = 1,2$. Therefore, we have to consider $\dfrac{\partial L(\bar{\underline{\alpha}})}{\partial \underline{\beta}}$ for these cases and to express them in terms of the true values of the thresholds, not the estimates $\bar{\underline{\alpha}}$. By Taylor Series Approximation, we get,

$$N^{-1/2} \frac{\partial L_{k,ij}(\bar{\underline{\alpha}}, \underline{\beta}_{k,ij})}{\partial \underline{\beta}_{k,ij}}$$

$$= N^{-1/2} \frac{\partial L_{k,ij}(\underline{\alpha}, \underline{\beta}_{k,ij})}{\partial \underline{\beta}_{k,ij}} + \underline{J}_{k,ij,ij} N^{1/2} (\bar{\underline{\alpha}} - \underline{\alpha}) + o_p(1), \tag{3.10}$$

where $\underline{J}_{k,ij,ij} = N^{-1} \dfrac{\partial^2 L_{k,ij}(\underline{\alpha},\underline{\beta}_{k,ij})}{\partial \underline{\beta}_{k,ij}\ \partial \underline{\alpha}}$ which converges in probability to a part of the information matrix corresponding to $(\underline{\alpha}', \underline{\beta}')'$.

It follows from (3.10), (3.9) can be written as

$$N^{1/2}\ (\hat{\underline{\beta}}(\bar{\underline{\alpha}}) - \underline{\beta}) = - \underline{I}^{-1} \left[N^{-1/2}\ \underline{\Delta}_{\underline{\beta}}(\underline{\alpha},\underline{\beta}) + \underline{J}\ N^{1/2}\ (\bar{\underline{\alpha}} - \underline{\alpha}) \right] + o_p(1), \tag{3.11}$$

where $\underline{J}$ is a matrix with blocks $\underline{J}_{k,ij,ij}$ for k = 1,2 and zero block matrix $\underline{0}$ when k = 3.

<u>Lemma 3.3</u> Under mild regularity conditions, the partial derivatives of $L_{k,ij}(\underline{\alpha},\underline{\beta}_{k,ij})$ with respect to $\underline{\beta}_{k,ij}$ satisfies

$$N^{-1/2} \frac{\partial L_{k,ij}(\underline{\alpha},\underline{\beta}_{k,ij})}{\partial \underline{\beta}_{k,ij}} = - \underline{I}_{k,ij,ij}\ N^{1/2}\ (\tilde{\underline{\beta}}_{k,ij} - \underline{\beta}_{k,ij}) + o_p(1)$$

where $\tilde{\underline{\beta}}_{k,ij}$ is the ML estimates of $\underline{\beta}_{k,ij}$ when k=1,2.

<u>Proof</u>

Assume that we know the true values of the thresholds, $\underline{\alpha}$, then we can have a system of likelihood equations,

$$\frac{\partial L_{k,ij}(\underline{\alpha},\tilde{\underline{\beta}}_{k,ij})}{\partial \underline{\beta}_{k,ij}} = \underline{0}$$

Then, by Taylor Series Approximation, we have,

$$N^{-1/2} \frac{\partial L_{k,ij}(\underline{\alpha}, \tilde{\underline{\beta}}_{k,ij})}{\partial \underline{\beta}_{k,ij}}$$

$$= N^{-1/2} \frac{\partial L_{k,ij}(\underline{\alpha}, \underline{\beta}_{k,ij})}{\partial \underline{\beta}_{k,ij}} + \left[N^{-1} \frac{\partial^2 L_{k,ij}(\underline{\alpha}, \underline{\beta}_{k,ij})}{\partial \underline{\beta}_{k,ij}\, \partial \underline{\beta}_{k,ij}} \right] \times$$

$$\left[N^{1/2} (\tilde{\underline{\beta}}_{k,ij} - \underline{\beta}_{k,ij}) \right] + o_p(N^{-1})$$

$$= N^{-1/2} \frac{\partial L_{k,ij}(\underline{\alpha}, \underline{\beta}_{k,ij})}{\partial \underline{\beta}_{k,ij}} + \underline{I}_{k,ij,ij}\, N^{1/2} (\tilde{\underline{\beta}}_{k,ij} - \underline{\beta}_{k,ij}) + o_p(1)$$

as $N^{-1} \frac{\partial^2 L_{k,ij}(\underline{\alpha}, \underline{\beta}_{k,ij})}{\partial \underline{\beta}_{k,ij}\, \partial \underline{\beta}_{k,ij}}$ converges in probability to the information matrix $\underline{I}_{k,ij,ij}$. Then, by rearranging the terms, Lemma 3.3 results.■

Thus, by Lemma 3.3, (3.10) can be expressed in the form

$$N^{-1/2}\, \bar{\underline{\Delta}}_{\underline{\beta}} = N^{-1/2} \frac{\partial L_{k,ij}(\bar{\underline{\alpha}}, \underline{\beta}_{k,ij})}{\partial \underline{\beta}_{k,ij}}$$

$$= - \underline{I}_{k,ij,ij}\, N^{1/2} (\tilde{\underline{\beta}}_{k,ij} - \underline{\beta}_{k,ij}) + \underline{J}_{k.ij,ij}\, N^{1/2} (\bar{\underline{\alpha}} - \underline{\alpha}) + o_p(1)$$

$$= A + B + o_p(1). \tag{3.12}$$

Pierce (1982) shows that $N^{1/2} (\tilde{\underline{\beta}}_{k,ij} - \underline{\beta}_{k,ij})$ and $N^{1/2} (\bar{\underline{\alpha}} - \underline{\alpha})$ are asymptotically independent, it follows that A and B are also asymptotically independent and hence, from (3.12) and Lemma 3.3, so do $N^{-1/2}\underline{\Delta}_{\underline{\beta}}$ and $N^{1/2} (\bar{\underline{\alpha}} - \underline{\alpha})$.

$\underline{\Delta}_{\underline{\beta}}$ consists functions $\frac{\partial L_{k,ij}(\underline{\alpha}, \underline{\beta})}{\partial \underline{\beta}_{k,ij}}$, k = 1,2,3, which are, in turn, functions of identical and independent random vectors, D_t, t = 1,...,N. Moreover, by the property of ML estimates that $E\left[\frac{\partial L_{k,ij}}{\partial \underline{\beta}_{k,ij}}\right] = 0$ for all k, we have $E(\underline{\Delta}_{\beta}) = 0$. Hence, by Central Limit Theorem,

$$N^{-1/2}\ \underline{\Delta}_{\beta} \stackrel{L}{===} N[\ \underline{0},\ \underline{\Omega}_{\beta}\ ]. \tag{3.13}$$

Together with (1) the estimates of the thresholds, $\bar{\underline{\alpha}}$, are asymptotically normal and (2) $N^{-1/2}\ \underline{\Delta}_{\beta}$ is independent of $N^{1/2}\ (\bar{\underline{\alpha}} - \underline{\alpha})$, we have the following result and theorem 3.1.

$$\begin{bmatrix} N^{-1/2}\ \underline{\Delta}_{\beta} \\ N^{1/2}\ (\bar{\underline{\alpha}} - \underline{\alpha}) \end{bmatrix} \stackrel{L}{===} N\left[\begin{pmatrix} \underline{0} \\ \underline{0} \end{pmatrix}, \begin{pmatrix} \underline{\Omega}_{\beta} & \underline{0} \\ \underline{0} & \underline{\Omega}_{\bar{\alpha}} \end{pmatrix} \right].$$

<u>Theorem 3.1</u> If $\hat{\underline{\beta}}(\bar{\underline{\alpha}})$ is the pseudo estimate of $\underline{\beta}$ and $\bar{\underline{\alpha}}$ is a consistent estimate of $\underline{\alpha}$, then, the asymptotic distribution of $N^{1/2}\ (\hat{\underline{\beta}}(\bar{\underline{\alpha}}) - \underline{\beta})$ is normal with zero mean vector and covariance matrix $\underline{\Omega}_{\beta(\bar{\alpha})}$ equals

$$\underline{I}^{-1}(\underline{\Omega}_{\beta} + \underline{J}\ \underline{\Omega}_{\bar{\alpha}}\ \underline{J}'\ )\underline{I}^{-1}.$$

<u>Proof</u>

Let $\underline{C} = (\ I^{*},\ I^{*})$, where I^{*} is the identity matrix with dimension equals that of $\underline{\Omega}_{\beta}$. Then, (3.11) can be expressed as

$$N^{1/2}\ (\ \hat{\underline{\beta}}(\bar{\underline{\alpha}}) - \underline{\beta}\ ) = -\ \underline{I}^{-1}\ \underline{C} \begin{pmatrix} N^{-1/2}\ \underline{\Delta}_{\beta} \\ \underline{J}\ N^{1/2}(\bar{\underline{\alpha}} - \underline{\alpha}) \end{pmatrix} + o_p(1).$$

Hence, by multivariate normal theory, we get,

$$N^{1/2}\ (\ \hat{\underline{\beta}}(\bar{\underline{\alpha}}) - \underline{\beta}\ ) \xrightarrow{L} N\left[\ \underline{0},\ \underline{I}^{-1}\ \underline{C}\begin{pmatrix} \underline{\Omega}_{\beta} & \underline{0} \\ \underline{0} & \underline{J}\ \underline{\Omega}_{\bar{\alpha}}\ \underline{J}' \end{pmatrix}\underline{C}'\ \underline{I}^{-1}\right],$$

or,

$$N^{1/2}\ (\ \hat{\underline{\beta}}(\bar{\underline{\alpha}}) - \underline{\beta}\ ) \xrightarrow{L} N\left[\ \underline{0},\ \underline{I}^{-1}(\ \underline{\Omega}_{\beta} + \underline{J}\ \underline{\Omega}_{\bar{\alpha}}\ \underline{J}'\ )\ \underline{I}^{-1}\right].\blacksquare$$

Let $\underline{T}$ be an appropriate selection matrix such that $\underline{T}\underline{\beta} = \underline{\sigma}$, then,

$$N^{1/2}\ (\bar{\underline{\sigma}} - \underline{\sigma}) \xrightarrow{L} N[\underline{0},\underline{\Gamma}],$$

where $\underline{\Gamma} = \underline{T}\ \underline{I}^{-1}(\ \underline{\Omega}_{\beta} + \underline{J}\ \underline{\Omega}_{\bar{\alpha}}\ \underline{J}'\ )\ \underline{I}^{-1}\ \underline{T}'$. Therefore, the joint asymptotic distribution of $\bar{\underline{\sigma}}$ has been established.

Stage III

The parameter vector $\underline{\theta}_o$ in the covariance structure $\underline{\Sigma}(\underline{\theta}_o)$ is estimated by minimizing the generalized least squares (GLS) function

$$Q(\underline{\theta}) = [\ \bar{\underline{\sigma}} - \underline{\sigma}(\underline{\theta})\ ]'\ \underline{W}^{-1}\ [\ \bar{\underline{\sigma}} - \underline{\sigma}(\underline{\theta})\ ],$$

where $\underline{W}$ is a positive definite matrix which converges in probability to $\underline{\Gamma}$.

Let $\hat{\underline{\theta}}$ be the GLS estimate of $\underline{\theta}_o$ by minimizing $Q(\underline{\theta})$. Follow the arguments given in Ferguson (1958), the following asymptotic properties of $\hat{\underline{\theta}}$ can be derived.

(1) $\hat{\underline{\theta}}$ is a consistent estimator of $\underline{\theta}_o$;

(2) The asymptotic distribution of $N^{1/2}\ (\ \hat{\underline{\theta}} - \underline{\theta}_o\ )$ is multivariate normal with zero mean vector and covariance matrix

$$(\partial \underline{\sigma}(\underline{\theta})/\partial \underline{\theta}) \ \underline{W}^{-1} \ (\partial \underline{\sigma}(\underline{\theta})/\partial \underline{\theta})';$$

(3) The asymptotic distribution of $NQ(\hat{\underline{\theta}})$ is chi-squared with degrees of freedom $p(p+1)/2-q$.

Based on the above results, some statistical inference can be performed such as (i) the chi-squared statistic $NQ(\hat{\underline{\theta}})$ in (3) can be used to test for the goodness-of-fit of the model $\underline{\Sigma} = \underline{\Sigma}(\underline{\theta})$, (ii) test of various hypothesis on $\underline{\theta}_o$ can be performed based on the result (2).

§3.3 Optimization Procedure and Simulation Study

§3.3.1 Optimization Procedure

The estimates of the thresholds and the polychoric correlations are obtained by following the Scoring procedure for polytomous-polytomous pair given in Chapter 2. Then, there exists an appropriate selection matrix $\underline{S}$ such that the weighting average of the estimates of the thresholds are picked out. In other words, $\underline{S}$ takes the values of $\underline{S}_{\underline{\beta}}$ in Chapter 2 except for the elements for picking the polychoric correlations which will be zero in $\underline{S}$. Hence, the consistent estimates $\bar{\underline{\alpha}}$ of the thresholds $\underline{\alpha}$ are obtained.

In stage II, the values of the thresholds, $\underline{\alpha}$, are fixed at $\bar{\underline{\alpha}}$. With slight modification of the procedures given in Chapter 2, we can minimize the negative log-likelihood functions by Scoring and Newton-Raphson algorithms. The difference is that we need to find the expressions of the first and second partial derivatives of $\Phi_2(\alpha,\beta;\underline{\mu},\underline{\Sigma})$ with respect to the parameters means, variances and covariance for implement the gradient vector and the Hessian or information matrix. Details for these expressions can be found in Poon & Lee, (1987).

The generalized least squares estimate $\hat{\underline{\theta}}$ of $\underline{\theta}$ is given by the q×1 vector which minimizes the GLS function $Q(\underline{\theta})$. Similarly, the minimization of GLS function cannot be solved in closed form, and hence, the Gauss-Newton algorithm (See, Lee & Jennrich(1979)) is used. Its basic step is given by

$$\Delta\underline{\theta} = - \zeta\ \underline{V}(\underline{\theta})^{-1}\ \dot{Q}(\underline{\theta})$$

where ζ is the step-size parameter described before and

$$\begin{aligned}\dot{Q}(\underline{\theta}) &= \partial Q(\underline{\theta})\ /\ \partial\underline{\theta} \\ &= (\partial\underline{\sigma}/\partial\underline{\theta}) \times (\partial Q(\underline{\theta})/\partial\underline{\sigma}) \\ &= 2\ (\ \partial\underline{\sigma}/\partial\underline{\theta}\ )\ \underline{W}^{-1}\ (\ \bar{\underline{\sigma}} - \underline{\sigma}\ ),\end{aligned}$$

and $\underline{V}(\underline{\theta}) = (\ \partial\underline{\sigma}/\partial\underline{\theta}\ )\ \underline{W}^{-1}\ (\ \partial\underline{\sigma}/\partial\underline{\theta}\ )'$.

The root mean squares $\|\dot{Q}(\underline{\theta})\|$ is used as the convergence criterion as before.

In order to obtain the expressions for $\dot{Q}(\underline{\theta})$ and $\underline{V}(\underline{\theta})$, we need to compute $\partial\underline{\sigma}/\partial\underline{\theta}$. But $\underline{\theta}$ is the population structural parameter vector which depends on the actual structure of the model. Hence, the expression for $\partial\underline{\sigma}/\partial\underline{\theta}$ is different for different structural models. Thus the expression will not be presented here.

Now, we turn to the approximation of the weight matrix $\underline{W}$ which is crucial in the generalized least squares estimation. As given in Theorem 3.1, we need consistent estimates of $\underline{I}$, $\underline{J}$, $\underline{\Omega}_{\underline{\beta}}$ and $\underline{\Omega}_{\bar{\underline{\alpha}}}$ to compute $\underline{W}$.

For $\underline{I}$, consistent estimates for each diagonal block matrix are needed which can be obtained by replacing the unknown parameters by its corresponding consistent estimates in (3.9). That is, the information matrix $\underline{I}_1(\hat{\underline{\beta}}_1)$ or the Hessian matrix $\underline{H}_k(\hat{\underline{\beta}}_k)$, k = 2,3, in the final iteration of the Scoring or Newton-Raphson algorithm for each sub-model is the consistent estimates for the corresponding block matrix of $\underline{I}$ which is denoted by $\hat{\underline{I}}$.

For $\underline{J}$, each of its block matrices can be estimated by replacing the unknown parameters by its consistent estimate in (3.10). However, the expression for the second partial derivative given is complicated and hence need a lot of computational time. Thus, the following natural consistent estimate for $\underline{J}$ is used. Let $\hat{\underline{J}}$ be the consistent estimate of $\underline{J}$, such that the r-th block of $\hat{\underline{J}}$ is given by $\hat{\underline{J}}_k$, k= 1,2:

$$\hat{\underline{J}}_{1,ij,ij} = N^{-1}\sum_{t=1}^{N}\left[\frac{\partial F_{1,ij,t}}{\partial \underline{\beta}_{1,ij}}\right]\left[\frac{\partial F_{1,ij,t}}{\partial \underline{\alpha}_{ij}}\right]'$$

$$\hat{\underline{J}}_{2,ij,ij} = N^{-1}\sum_{t=1}^{N}\left[\frac{\partial F_{2,ij,t}}{\partial \underline{\beta}_{2,ij}}\right]\left[\frac{\partial F_{2,ij,t}}{\partial \underline{\alpha}_{i}}\right]'$$

$$\text{where}\begin{cases} i,j=1,\ldots,p_1, i>j;\; r = (i-2)(i-1)/2+j & \text{for } k = 1 \\ i=1,\ldots,p_1; j=1,\ldots,p_2;\; r = (p_1-1)p_1/2+(i-1)p_1+j & \text{for } k = 2. \end{cases}$$

For $\underline{\Omega}_{\underline{\alpha}}$, a consistent estimate of $\underline{\Omega}_{\underline{\alpha}}$ can be obtained by referring to Lee, Poon & Bentler (1990b). However, their estimate needs a lot of computer storage, hence, in this thesis, another approximation is used. Recall the notations and equation (2.15) in Chapter 2, there exists a selection matrix $\underline{S}_{\alpha}$ such that $\underline{S}_{\alpha}\underline{\gamma} = \underline{\alpha}$. Now, in order to obtain a consistent estimate for $\underline{\Omega}_{\underline{\alpha}}$,

$r = (p_1-1)p_1/2 + (i-1)p_2 + j;\ s = (m-2)(m-1)/2 + n;$

(3) for $k = 3$ and $i = 1$,

$i, j = 1,\ldots,p_2;\ i > j;\ m, n = 1,\ldots,p_1;\ m > n;$

$r = (p_1-1)p_1/2 + p_1p_2 + (i-2)(i-1)/2 + j;\ s = (m-2)(m-1)/2 + n;$

(4) for $k = 2$ and $l = 2$,

$i, m = 1,\ldots,p_1;\ j, n = 1,\ldots,p_2;$

$r = (p_1-1)p_1/2 + (i-1)p_2 + j;\ s = (p_1-1)p_1/2 + (m-1)p_2 + n;$

(5) for $k = 3$ and $l = 2$,

$i, j = 1,\ldots,p_2;\ i > j;\ m = 1,\ldots,p_1;\ n = 1,\ldots,p_2;$

$r = (p_1-1)p_1/2 + p_1p_2 + (i-2)(i-1)/2 + j;$

$s = (p_1-1)p_1/2 + (m-1)p_2 + n;$

(6) for $k = 3$ and $l = 3$,

$i, j, m, n = 1,\ldots,p_2;\ i > j;\ m > n;$

$r = (p_1-1)p_1/2 + p_1p_2 + (i-2)(i-1)/2 + j;$

$s = (p_1-1)p_1/2 + p_1p_2 + (m-2)(m-1)/2 + n;$

and other off-diagonal elements can be obtained by the symmetry of $\underline{\Omega}_{\beta}$.

Thus, a good choice of $\underline{W}$ is given by

$$\underline{T}\ \hat{\underline{I}}^{-1}\ (\ \hat{\underline{\Omega}}_{\beta} + \hat{\underline{J}}\ \hat{\underline{\Omega}}_{\alpha}\ \hat{\underline{J}}'\ )\ \hat{\underline{I}}^{-1}\ \underline{T}'.$$

§3.3.2 Monte Carlo Study

Based on the theories developed in Section 2 and the minimization procedure set up in Section 3.3.1, a computer program written in FORTRAN IV with double precision has been implemented. A Monte Carlo study was conducted to demonstrate the theories established. Exact continuous data are simulated from a multivariate normal distribution with zero mean vector we need to find consistent estimates for $\underline{K}$ and $\underline{\nabla}_{\underline{\gamma}}$. A consistent estimate $\hat{\underline{K}}$ of $\underline{K}$ can be easily found by the same method for that of $\underline{I}$ described before. Let $\hat{\underline{\nabla}}_{\underline{\gamma}}$ be the consistent estimate of $\underline{\nabla}_{\underline{\gamma}}$ such that the (r,s)th block of this consistent estimate is given by

$$\hat{\underline{\nabla}}_{\underline{\gamma}_{ij,mn}} = N^{-1} \sum_{t=1}^{N} \left[\frac{\partial F_{1,ij,t}}{\partial \underline{\alpha}_{ij}} \right] \left[\frac{\partial F_{1,mn,t}}{\partial \underline{\alpha}_{mn}} \right]',$$

where i, j, m, n = $1,\ldots,p_1$, i>j; m>n; r=(i-2)(i-1)/2+j; s=(m-2)(m-1)/2+n. Then, a consistent estimate $\hat{\underline{\Omega}}_{\bar{\underline{\alpha}}}$ for the asymptotic covariance matrix $\underline{\Omega}_{\bar{\underline{\alpha}}}$ for the PML estimate $\bar{\underline{\alpha}}$ is given by

$$\underline{S}_{\underline{\alpha}} \hat{\underline{K}}^{-1} \hat{\underline{\nabla}}_{\underline{\gamma}} \hat{\underline{K}}^{-1} \underline{S}_{\underline{\alpha}}'.$$

Similarly for $\underline{\Omega}_{\underline{\beta}}$, the (r,s)th block of the consistent estimate $\hat{\underline{\Omega}}_{\underline{\beta}}$ is given by

$$\hat{\underline{\Omega}}_{\underline{\beta}_{ij,mn}} = N^{-1} \sum_{t=1}^{N} \left[\frac{\partial F_{k,ij,t}}{\partial \underline{\beta}_{k,ij}} \right] \left[\frac{\partial F_{1,mn,t}}{\partial \underline{\beta}_{1,mn}} \right]',$$

where

(1) for k = 1 and l = 1,

i, j, m, n = $1,\ldots,p_1$; i > j; m > n;

r = (i-2) (i-1)/2 + j; s = (m-2) (m-1)/2 + n;

(2) for k = 2 and l = 1,

i = $1,\ldots,p_1$; j = $1,\ldots,p_2$; m, n = $1,\ldots,p_1$; m > n;

and covariance matrix $\underline{\Sigma}$, which was supposed to follow a confirmatory factor analysis model (Lawley & Maxwell, 1971), that is,

$$\underline{\Sigma} = \underline{F}\ \underline{M}\ \underline{F}' + \underline{E},$$

where $\underline{F}$ is the factor loading matrix, $\underline{M}$ and $\underline{E}$ are covariance matrices of the factors and residuals respectively. There are 6 polytomous variables each with 3 categories and 2 interval variables. The simulated random vector was transformed to the corresponding polytomous variables Z_is by using the pre-assigned thresholds,

$$\underline{\alpha}_i = (\ -\infty,\ -0.5,\ 0.5,\ \infty\ ),\ i = 1,\ldots,6;$$

and the interval data set $(w_{t,1}, w_{t,2})$ was obtained by adding and subtracting a small values, say c_{1t} and c_{2t}, from the exact continuous data. If N is the sample size of a data set considered, the corresponding values c_{1t} and c_{2t} are given by (0.01, 0.03) for t= 1,...,N/4; (0.02, 0.04) for t= N/4 + 1,...,3N/4; and (0.03, 0.01) for t = 3N/4 + 1,..., N.

The population values of parameter matrices are given by

$$\underline{F}' = \begin{bmatrix} 0.8 & 0.8 & 0.8 & 0.8 & 0.0^* & 0.0^* & 0.0^* & 0.0^* \\ 0.0^* & 0.0^* & 0.0^* & 0.0^* & 0.8 & 0.8 & 0.8 & 0.8 \end{bmatrix},$$

$$\underline{M} = \begin{bmatrix} 1.0^* & 0.5 \\ 0.5 & 1.0^* \end{bmatrix},$$

and

$$\underline{E} = \text{diag}\ [\ 0.36\ \ 0.36\ \ 0.36\ \ 0.36\ \ 0.36\ \ 0.36\ \ 0.36\ \ 0.36\ ],$$

where the off diagonal elements of $\underline{E}$ and the values with an asterisk are treated as fixed parameters and were not estimated.

Estimates for the structural parameter $\underline{\theta}_o$ based on various sample sizes, N = 300, 400, 500, 700 and 1000, are obtained by the written program and the results are presented in Tables 2. 40 replications have been performed for each sample size. Similar to the simulation study in Chapter 2, the five statistics are reported. Besides, another index, maximum difference between the estimates of the standard errors and the sample standard errors in the 40 replications, MaxD, and the p-values for the Kolmogorov-Smirnov statistic for testing the hypothesis H_o that $NQ(\hat{\underline{\theta}})$ is chi-squared distributed are also reported.

It can be observed that

(1) the estimates for the structural parameters are pretty good;

(2) in many cases, the root mean squares errors decrease as N increases;

(3) the p-values for the hypothesis H_o are not significant and hence H_o is not rejected at the 0.05 significant level for all N being studied;

(4) the estimates of standard errors are not very satisfactory. For example, we always overestimate the standard errors for the covariances of the residuals which are corresponding to the polytomous variables and underestimate that for the factor loading corresponding to the interval variables;

(5) the MaxD decreases as N increases.

The result may be due to the facts that the sample size is not large enough in our cases as we have use two asymptotic properties of the estimates. Besides, the number of replications may also be too small to reflect the properties of our estimates. Hence, in order to obtain estimates of the structural parameters by using the method developed above, a very large data set is needed. However, in some cases, it might be hard to collect such a large data set in real-life situation.

Chapter 4

Two-stage Procedure for Correlation Structure Analysis

In Chapter 3, a three-stage procedure is employed to analyze the covariance structure model. It is well known that no variance information is available for polytomous variables. Therefore, to identify the model, the thresholds are first estimated by fixing the variances of the polytomous variables at 1, and then the variances are estimated by fixing the thresholds at the estimated values. As a result, the estimates of the variances of the polytomous variables are always very close to 1. That is, we are essentially considering the correlation structure of the polytomous variables. For the interval variables, information for variances is available and hence we can estimate the covariance structure. Thus, a mixture of correlations and covariances is used. However, in some situations, for example, in a scale invariant model, the use of correlations only may be enough. Therefore, it is worthwhile to consider a simpler two-stage method to handle the correlation structure. In this chapter, a two-stage procedure is developed for analyzing the correlation structure model. The outline of the two-stage procedure is presented in Section 2. The asymptotic statistical properties of the estimates obtained is studied via a Monte Carlo study in Section 3. Finally, a comparison is conducted to investigate the performances of these two procedures with same correlation structure in Section 4.

§4.1 Model

Similar to Chapter 2, the general multivariate correlation structure model with polytomous and interval variables is considered. The two latent continuous variables, $p_1 \times 1$, $\underline{X}$ and, $p_2 \times 1$, $\underline{Y}$ are multivariate normal distributed with mean vector $(\underline{0}', \underline{\mu}_y')'$ and covariance matrix,

$$\underline{\Sigma} = \begin{bmatrix} \underline{R}_{xx} & \underline{C}_{xy} \\ \underline{C}_{yx} & \underline{\Sigma}_{yy} \end{bmatrix}.$$

Let $\underline{R}$ be the corresponding correlation matrix and has the form

$$\underline{R} = \begin{bmatrix} \underline{R}_{xx} & \underline{R}_{xy} \\ \underline{R}_{yx} & \underline{R}_{yy} \end{bmatrix}.$$

Now, we consider the correlation instead of the covariance structure, that is, $\underline{R} = \underline{R}(\underline{\theta}_o)$ where $\underline{\theta}_o$ is a $q \times 1$ unknown structural parameter vector. Moreover, $\underline{R}$ is also assumed to be identified. The corresponding polytomous and interval variables $\underline{Z}$ and $\underline{W}_1, \ldots, \underline{W}_{p_2}$ given in Chapter 2 are also assumed to be observed. The relationship between $\underline{X}$ and $\underline{Z}$ is given by (2.1) and the corresponding interval observation is $(w_{j,1}, w_{j,2})$ with $w_{j,1} < w_{j,2}$ for $j = 1, \ldots, p_2$.

§4.2 Two-stage Estimation Method

The data sample D_t is supposed to be obtained for $t = 1, \ldots, N$. We also define $\alpha_{t,i,k_t(i)} = \alpha_{i,k(i)}$ if $k_t(i) = k(i)$ for $i = 1, \ldots, p_1$.

Similar to Chapter 3, only the partition maximum likelihood estimation is considered. In stage I, the estimates of the covariance matrix $\underline{\Sigma}$, and hence then that of the correlation matrix $\underline{R}$, without imposing any structure is obtained. The parameter vector corresponding to the covariance matrix $\underline{\Sigma}$ is taken as $\underline{\beta}$ which consists of the thresholds of the polytomous variables $\underline{\alpha}$, the means of the interval variables $\underline{\mu}_y$, and the polychoric correlations, covariances and variances in $\underline{\Sigma}$, $\underline{\sigma} = \{\rho_{ij},\ i,j = 1,\ldots,p_1;\ i>j;\ \sigma_{ij},\ i = 1,\ldots,p_1,\ j = p_1+1,\ldots,p;\ \sigma_{ij},\ i,\ j = p_1+1,\ldots,p;\ i\geq j\}$. That is, $\underline{\beta} = (\underline{\alpha}', \underline{\mu}_y', \underline{\sigma}')'$. In stage II, the structural parameter vector $\underline{\theta}_o$ will be estimated by the generalized least squares method.

Stage I

Follow the discussion and notations given in §2.3.1 to §2.3.4, the normality of partition maximum likelihood estimator $\bar{\underline{\sigma}}$ for $\underline{\sigma}$ can be established. In the following, the normality of the estimate of the lower triangular elements of $\underline{R}$ is derived. Recall the notations and the result (2.15) obtained in Chapter 2. Namely, let

$$\underline{\gamma} = (\underline{\gamma}_{1,21}',\ldots,\underline{\gamma}_{1,p_1(p_1-1)}',\underline{\gamma}_{2,11}',\ldots,\underline{\gamma}_{2,p_1p_2}',\underline{\gamma}_{3,21}',\ldots,\underline{\gamma}_{3,p_2(p_2-1)}')'$$

and

$$\underline{\Delta}_{\gamma} = \frac{\partial L}{\partial \underline{\gamma}} = \left[\frac{\partial L_{1,21}'}{\partial \underline{\gamma}_{1,21}}, \ldots, \frac{\partial L_{1,p_1(p_1-1)}'}{\partial \underline{\gamma}_{1,p_1(p_1-1)}}, \frac{\partial L_{2,11}'}{\partial \underline{\gamma}_{2,11}}, \ldots, \frac{\partial L_{2,p_1p_2}'}{\partial \underline{\gamma}_{2,p_1p_2}}, \frac{\partial L_{3,21}'}{\partial \underline{\gamma}_{3,21}}, \ldots, \frac{\partial L_{3,p_2(p_2-1)}'}{\partial \underline{\gamma}_{3,p_2(p_2-1)}} \right]', \tag{4.1}$$

then, the following result is obtained. That is,

$$N^{1/2} (\hat{\underline{\gamma}} - \underline{\gamma}) \xrightarrow{L} N[\ \underline{0},\ \underline{K}^{-1}\ \underline{\nabla}_{\gamma}\ \underline{K}^{-1}\ ], \tag{4.2}$$

where $\underline{K}$ is the information matrix of $\underline{\gamma}$ which is a diagonal block matrix of the form given in Chapter 2, and $\underline{V}_{\underline{\gamma}}$ is the asymptotic covariance matrix of $\partial L/\partial\underline{\gamma}$.

Now, there exists a selection matrix $\underline{T}$ which picks out $\underline{\sigma}$ from $\underline{\gamma}$, that is, $\underline{T}\ \underline{\gamma} = \underline{\sigma}$. Then, by the multivariate normal theory, we have the following result.

$$N^{1/2}\ (\ \bar{\underline{\sigma}} - \underline{\sigma}\ ) \xrightarrow{L} N[\underline{0},\ \underline{T}\ \underline{K}^{-1}\ \underline{V}_{\underline{\gamma}}\ \underline{K}^{-1}\ \underline{T}']\ .$$

Let $\underline{\rho} = (\ \rho_{ij},\ i,\ j = 1,\ldots,p;\ i>j\ )$ be the vector which consists of all the non-duplicated off-diagonal elements of $\underline{R}$. Then, the PML estimate $\bar{\underline{\rho}}$ of $\underline{\rho}$ can be obtained by normalization of $\bar{\underline{\sigma}}$. Moreover, under mild regularity conditions, by Delta Theorem, the PML estimate $\bar{\underline{\rho}}$ of $\underline{\rho}$ follows a multivariate normal distribution,

$$N^{1/2}\ (\ \bar{\underline{\rho}} - \underline{\rho}\ ) \xrightarrow{L} N[\underline{0},\ \underline{\Gamma}],$$

where $\underline{\Gamma} = \dfrac{\partial\underline{\rho}}{\partial\underline{\sigma}}\ \underline{T}\ \underline{K}^{-1}\ \underline{V}_{\underline{\gamma}}\ \underline{K}^{-1}\ \underline{T}'\ \dfrac{\partial\underline{\rho}}{\partial\underline{\sigma}}'$.

Stage II

The GLS estimate $\hat{\underline{\theta}}$ of $\underline{\theta}_o$ is obtained by minimizing the GLS function $Q(\underline{\theta})$,

$$Q(\underline{\theta}) = [\ \bar{\underline{\rho}} - \underline{\rho}(\underline{\theta})\ ]'\ \underline{W}^{-1}\ [\ \bar{\underline{\rho}} - \underline{\rho}(\underline{\theta})\ ],$$

where $\underline{W}$ is a positive definite matrix which converges in probability to $\underline{\Gamma}$. Follow the arguments given in Ferguson(1958), the GLS estimate $\hat{\underline{\theta}}$ shares the nice properties (1) to (3) given in Chapter 3, namely,

(1) $\hat{\underline{\theta}}$ is consistent;

(2) $N^{1/2}$ $(\hat{\underline{\theta}} - \underline{\theta})$ is multivariate normal with mean vector $\underline{0}$ and covariance matrix $(\partial\underline{\rho}(\underline{\theta})/\partial\underline{\theta})\ \underline{W}^{-1}\ (\partial\underline{\rho}(\underline{\theta})/\partial\underline{\theta})'$;

(3) $NQ(\hat{\underline{\theta}})$ is asymptotically chi-squared distributed with degrees of freedom $p(p-1)/2 - q$.

Hence, the statistical inference discussed in Chapter 3 can also be performed in this case.

§4.3 Optimization Procedure and Monte Carlo Study

§4.3.1 Optimization Procedure

The estimates of the covariance matrix $\underline{\Sigma}$ and hence the estimates of the correlation matrix $\underline{R}$ can be obtained by the procedure of using the Scoring and Newton-Raphson algorithms in Chapter 2. Now, we only need to find a consistent estimate of $\underline{\Gamma}$, which can be estimated by consistent estimates of $\underline{K}$ and $\underline{\nabla}_{\underline{\gamma}}$ as the exact expressions for $\partial\underline{\rho}/\partial\underline{\sigma}$ can be easily derived from basic matrix calculus (see, e.g. Bentler & Lee, 1978).

Similar to Chapter 3, a consistent estimate $\hat{\underline{K}}$ of $\underline{K}$ can be obtained by the estimated information matrix or the Hessian matrix obtained in the final iteration of the Scoring or the Newton-Raphson algorithm.

For $\underline{\nabla}_{\underline{\gamma}}$, the (r,s)th block of a consistent estimate $\hat{\underline{\nabla}}_{\underline{\gamma}}$ is given by

$$\hat{\underline{\nabla}}_{\underline{\gamma} ij,mn} = N^{-1} \sum_{t=1}^{N} \left[\frac{\partial F_{k,ij,t}}{\partial \underline{\gamma}_{k,ij}} \right] \left[\frac{\partial F_{l,mn,t}}{\partial \underline{\gamma}_{l,mn}} \right]',$$

where

(1) for $k = 1$ and $l = 1$,

$i, j, m, n = 1,\ldots,p_1$; $i > j$; $m > n$;

$r = (i-2)(i-1)/2 + j$; $s = (m-2)(m-1)/2 + n$;

(2) for $k = 2$ and $l = 1$,

$i = 1,\ldots,p_1$; $j = 1,\ldots,p_2$; $m, n = 1,\ldots,p_1$; $m > n$;

$r = (p_1-1)p_1/2 + (i-1)p_2 + j$; $s = (m-2)(m-1)/2 + n$;

(3) for $k = 3$ and $l = 1$,

$i, j = 1,\ldots,p_2$; $i > j$; $m, n = 1,\ldots,p_1$; $m > n$;

$r = (p_1-1)p_1/2 + p_1p_2 + (i-2)(i-1)/2 + j$; $s = (m-2)(m-1)/2 + n$;

(4) for $k = 2$ and $l = 2$,

$i, m = 1,\ldots,p_1$; $j, n = 1,\ldots,p_2$;

$r = (p_1-1)p_1/2 + (i-1)p_2 + j$; $s = (p_1-1)p_1/2 + (m-1)p_2 + n$;

(5) for $k = 3$ and $l = 2$,

$i, j = 1,\ldots,p_2$; $i > j$; $m = 1,\ldots,p_1$; $n = 1,\ldots,p_2$;

$r = (p_1-1)p_1/2 + p_1p_2 + (i-2)(i-1)/2 + j$;

$s = (p_1-1)p_1/2 + (m-1)p_2 + j$;

(6) for $k = 3$ and $l = 3$,

$i, j, m, n = 1,\ldots,p_2$; $i > j$; $m > n$;

$r = (p_1-1)p_1/2 + p_1p_2 + (i-2)(i-1)/2 + j$;

$s = (p_1-1)p_1/2 + p_1p_2 + (m-2)(m-1)/2 + n$;

and $F_{k,ij,t}$ are the functions given in Chapter 2. Thus, a good choice of $\underline{W}$ is given by

$$(\partial \underline{\rho}/\partial \underline{\sigma})\ \underline{T}\ \hat{\underline{K}}^{-1}\ \hat{\underline{\nabla}}_{\underline{\gamma}}\ \hat{\underline{K}}^{-1}\ \underline{T}'\ (\partial \underline{\rho}/\partial \underline{\sigma})'.$$

§4.3.2 Monte Carlo Study

Based on the theories developed, a computer program written in FORTRAN IV with double precision has been implemented to obtain the two-stage estimates. Similar to Chapter 3, a Monte Carlo simulation is carried out to study the performance of our estimates.

Exact continuous data coming from a multivariate normal distribution with zero mean vector and covariance matrix $\underline{\Sigma}$ are simulated. The population covariance matrix $\underline{\Sigma}$ is assumed equal to the correlation matrix $\underline{R}$ which follows the structure given by

$$\underline{R} = \underline{F}\,\underline{M}\,\underline{F}' + \underline{E},$$

where the factor loading matrix $\underline{F}$ has the form,

$$\underline{F}' = \begin{bmatrix} 0.8 & 0.8 & 0.8 & 0.8 & 0.0^* & 0.0^* & 0.0^* & 0.0^* \\ 0.0^* & 0.0^* & 0.0^* & 0.0^* & 0.8 & 0.8 & 0.8 & 0.8 \end{bmatrix};$$

and the covariance matrix of the factor $\underline{M}$ has the form,

$$\underline{M} = \begin{bmatrix} 1.0^* & 0.5 \\ 0.5 & 1.0^* \end{bmatrix};$$

and the covariance matrix of the residuals $\underline{E}$ satisfies the following constraints,

(1) $\underline{E}$ is diagonal;

(2) the diagonal elements of $\underline{R}$ equal to one;

and the values with * are fixed parameters and were not estimated. Therefore, there are totally 9 free parameters as the diagonal elements in $\underline{E}$

are determined by the constraint (2). Then, the data are transformed to the corresponding polytomous and interval observations according to the rules given in Chapter 3.

Estimates of the structural parameter $\underline{\theta}_o$ can be obtained by the written program and a simulation is conducted with the above population values and various sample sizes. The results are presented in Tables 3. Similarly, 40 replications are performed and the six statistics and the index MaxD are reported. The following phenomenons are observed.

(1) the estimates for the structural parameters are very close to the true values;

(2) the root mean square errors decreases with the sample size N;

(3) the p-values for the K-S statistic for testing the hypothesis H_o given in Chapter 3 indicate that H_o is not rejected at $\alpha = 0.05$ in all cases except for N = 300. From Fig. 1-5, the observed statistics deviate from chi square even when the sample size is 700. However, improvement is shown and the observed statistics become chi-square distributed when N = 1000. Hence the result agrees with the theoretical assertion that $NQ(\hat{\underline{\theta}})$ is asymptotic chi-squared distributed;

(4) the averages of the estimated standard errors are close to the sample standard errors;

(5) the MaxD decreases with sample size N.

Hence, we can conclude that our estimates behave well for various sample sizes.

§4.4 Comparison of Two Methods

If the general covariance matrix considered in Chapter 3 is in fact a correlation matrix, then, both the three-stage procedure and the two-stage procedure discussed are applicable. So, it is of interest to study the performances of these two procedures when the same correlation structure model is under consideration. Several examples are presented to compare the two procedures.

Exact continuous multivariate normal variates are generated with zero mean vector and the correlation matrix $\underline{R}$, with the structure,

$$\underline{R} = \underline{F}\ \underline{M}\ \underline{F}' + \underline{E},$$

where $\underline{F}$ is the factor loading matrix, $\underline{M}$ and $\underline{E}$ are the covariance matrices of the factors and the residuals respectively. The population structure of $\underline{F}$, $\underline{M}$, and $\underline{E}$ are taken as,

$$\underline{F}' = \begin{bmatrix} 0.8 & 0.8 & 0.8 & 0.8 & 0.0^* & 0.0^* & 0.0^* & 0.0^* \\ 0.0^* & 0.0^* & 0.0^* & 0.0^* & 0.8 & 0.8 & 0.8 & 0.8 \end{bmatrix},$$

$$\underline{M} = \begin{bmatrix} 1.0^* & 0.5 \\ 0.5 & 1.0^* \end{bmatrix},$$

and $\underline{E} = \text{diag}\ [\ 0.36\ \ 0.36\ \ 0.36\ \ 0.36\ \ 0.36\ \ 0.36\ \ 0.36\ \ 0.36\ ].$

In both procedures, the values with * and the off-diagonal elements of $\underline{E}$, are treated as fixed parameters and will not be estimated. However, the diagonal elements in $\underline{E}$ are treated differently for the two procedures. They are considered as free parameters in the three-stage procedure while as dependent parameters in the two-stage procedure. The generated data are transformed to the corresponding polytomous and interval observations according to the criteria in Chapter 3.

Samples with sizes 100, 400, 500 and 1000 are simulated. Then, the estimates of the structural parameter vector are obtained by the two methods with same stopping criterion. Each PML estimate and its corresponding standard error are reported and the result is presented in Tables 4.

It is observed that the two methods give similar estimates and perform equally well, especially when the sample size is large. Besides, the model is not rejected in all cases as the observed χ^2 values are smaller than the upper 0.05 percentile of chi-squared distribution with degrees of freedom 19. Moreover, the observed chi-squared values are close for both procedures. However, when the sample size is small, say, N = 100, the PML estimates for the factor loadings corresponding to the interval variates differ quite a lot from the true values for the three-stage procedure while these are still very close to the true values for the two-stage procedure. Moreover, the PML estimates of some elements in the covariance matrix of the residual, $\underline{E}$, deviate from the true values for the three-stage procedure. From these examples, the two-stage procedure is recommended when correlation structure is considered for small samples. Of course, in order to draw more reliable conclusions, further simulation studies are required.

Chapter 5

Conclusion

In this thesis, the maximum likelihood estimate for the correlations between the polytomous and interval variables are obtained via the partition maximum likelihood estimation proposed by Poon & Lee (1987). It follows from the statistical theories that this maximum likelihood estimate has nice asymptotic properties. Based on the simulation study, we observe that the estimates are very accurate in various conditions. Moreover, a computationally efficient three-stage estimation procedure has been established for analyzing structural equation models when these two kinds of data are involved. In the first stage, the partition ML estimate of the thresholds of the polytomous variables are given. It is shown that this partition ML estimates are asymptotically jointly multivariate normal. In the second stage, the pseudo ML estimate of the covariance matrix is produced by fixing the thresholds at the values given in stage one. Asymptotic distribution of the pseudo ML estimates is derived and is jointly multivariate normal. With the normality of the estimates, the generalized least squares method is employed to estimate the structural parameters. Nice asymptotic properties of the GLS estimates for statistical inference of the model have also been provided. However, we observe from our simulation study that the method tends to overestimate the standard errors of the residuals estimates, especially when the sample size is small. Hence, misleading conclusion may be resulted when making inferences on the residual covariance matrix. Therefore, this method should be used only when the sample size is large enough.

A two-stage procedure for the analysis of correlation structure model has been developed. The first stage estimates the correlation matrix via the partition ML method. It is shown that the asymptotic distribution of this estimate is multivariate normal. Making use of this result, the second stage employed the generalized least squares estimation to obtain the estimate of the structural parameters in a correlation structure. Basic statistical properties of the estimates have been derived which can be used to perform various statistical inferences. Based on the simulation result, the estimates for the parameters and its standard errors are very accurate even when the sample size is small.

Several examples are given to illustrate the performances of the two procedures when a same correlation structure is considered. These examples indicate that the two methods give close estimates for the parameters under large samples while the two-stage procedure produces more accurate estimates under small samples.

Clearly, there are still a lot of practical problems that needed to be studied, such as the presence of missing observations in either or both kinds of variables and the involvement of the multisamples. These provide important research topics in the future.

Table 1a(i).1

Simulation Result of Threshold and Correlation Estimation

Symmetric distribution of Z_1 and Z_2 with small length for Y_1 and Y_2

(N = 70)

Parameter	Mean	RMS	$\overline{\text{S.E.}}$	S.E.	Ratio
μ_1 =0.0	0.0091	0.1189	0.1200	0.1204	0.9969
μ_2 =0.0	-0.0010	0.1359	0.1376	0.1188	1.1583
σ_{33}=1.0	1.0047	0.1844	0.1867	0.1700	1.0983
σ_{44}=1.0	1.0054	0.2063	0.2088	0.1702	1.2272
$\alpha_{1,2}$=-0.5	-0.5090	0.1590	0.1607	0.1560	1.0303
$\alpha_{1,3}$=0.5	0.5256	0.1431	0.1426	0.1460	0.9772
$\alpha_{2,2}$=-0.5	-0.5178	0.1490	0.1498	0.1561	0.9601
$\alpha_{2,3}$=0.5	0.4949	0.1397	0.1414	0.1453	0.9733
ρ_{21}=0.5	0.5153	0.1030	0.1031	0.1152	0.8957
ρ_{31}=0.5	0.5045	0.0982	0.0994	0.1016	0.9783
ρ_{32}=0.5	0.4880	0.0913	0.0916	0.1042	0.8795
ρ_{41}=0.5	0.5145	0.0829	0.0827	0.1019	0.8114
ρ_{42}=0.5	0.5072	0.1125	0.1138	0.1021	1.1143
σ_{43}=0.5	0.4998	0.1554	0.1573	0.0884	1.7800

Table 1a(i).2

Simulation Result of Threshold and Correlation Estimation

Symmetric distribution of Z_1 and Z_2 with small length for Y_1 and Y_2

(N = 100)

Parameter	Mean	RMS	$\overline{\text{S.E.}}$	S.E.	Ratio
μ_1 =0.0	0.0093	0.0802	0.0807	0.1008	0.8009
μ_2 =0.0	-0.0008	0.0918	0.0930	0.0988	0.9408
σ_{33}=1.0	1.0129	0.1443	0.1455	0.1434	1.0151
σ_{44}=1.0	0.9931	0.1687	0.1707	0.1406	1.2143
$\alpha_{1,2}$=-0.5	-0.4918	0.1123	0.1134	0.1291	0.8780
$\alpha_{1,3}$=0.5	0.4799	0.1277	0.1277	0.1215	1.0508
$\alpha_{2,2}$=-0.5	-0.4921	0.1270	0.1284	0.1298	0.9893
$\alpha_{2,3}$=0.5	0.4850	0.1221	0.1456	0.1216	1.1974
ρ_{21}=0.5	0.5144	0.0974	0.0976	0.0966	1.0106
ρ_{31}=0.5	0.5257	0.1116	0.1100	0.0826	1.3322
ρ_{32}=0.5	0.5096	0.1076	0.1085	0.0844	1.2855
ρ_{41}=0.5	0.5306	0.0860	0.0814	0.0824	0.9884
ρ_{42}=0.5	0.5007	0.1178	0.1193	0.0849	1.4048
σ_{43}=0.5	0.5023	0.1244	0.1260	0.0738	1.7070

Table 1a(i).3

Simulation Result of Threshold and Correlation Estimation

Symmetric distribution of Z_1 and Z_2 with small length for Y_1 and Y_2

(N = 200)

Parameter	Mean	RMS	$\overline{\text{S.E.}}$	S.E.	Ratio
μ_1 =0.0	0.0113	0.0734	0.0734	0.0705	1.0423
μ_2 =0.0	0.0143	0.0679	0.0673	0.0703	0.9568
σ_{33}=1.0	0.9974	0.1084	0.1098	0.0997	1.1009
σ_{44}=1.0	0.9938	0.1023	0.1034	0.0994	1.0402
$\alpha_{1,2}$=-0.5	-0.5045	0.0943	0.0954	0.0924	1.0323
$\alpha_{1,3}$=0.5	0.4917	0.0827	0.0833	0.0852	0.9785
$\alpha_{2,2}$=-0.5	-0.5046	0.0830	0.0840	0.0921	0.9121
$\alpha_{2,3}$=0.5	0.5033	0.0726	0.0734	0.0852	0.8615
ρ_{21}=0.5	0.4921	0.0701	0.0705	0.0700	1.0076
ρ_{31}=0.5	0.4914	0.0626	0.0628	0.0614	1.0223
ρ_{32}=0.5	0.4827	0.0622	0.0605	0.0622	0.9736
ρ_{41}=0.5	0.4837	0.0507	0.0486	0.0622	0.7818
ρ_{42}=0.5	0.4940	0.0622	0.0627	0.0614	1.0219
σ_{43}=0.5	0.4854	0.0873	0.0871	0.0531	1.6396

Table 1a(i).4

Simulation Result of Threshold and Correlation Estimation

Symmetric distribution of Z_1 and Z_2 with small length for Y_1 and Y_2

(N = 300)

Parameter	Mean	RMS	$\overline{\text{S.E.}}$	S.E.	Ratio
μ_1 =0.0	0.0110	0.0468	0.0460	0.0583	0.7894
μ_2 =0.0	0.0071	0.0546	0.0548	0.0571	0.9601
σ_{33}=1.0	1.0262	0.0951	0.0926	0.0838	1.1048
σ_{44}=1.0	0.9869	0.0791	0.0791	0.0806	0.9812
$\alpha_{1,2}$=-0.5	-0.4987	0.0771	0.0780	0.0750	1.0400
$\alpha_{1,3}$=0.5	0.4910	0.0718	0.0721	0.0695	1.0367
$\alpha_{2,2}$=-0.5	-0.4866	0.0756	0.0754	0.0751	1.0045
$\alpha_{2,3}$=0.5	0.5092	0.0823	0.0829	0.0696	1.1899
ρ_{21}=0.5	0.4940	0.0571	0.0575	0.0572	1.0067
ρ_{31}=0.5	0.4848	0.0561	0.0547	0.0506	1.0823
ρ_{32}=0.5	0.5177	0.0466	0.0437	0.0486	0.8982
ρ_{41}=0.5	0.5028	0.0556	0.0562	0.0495	1.1354
ρ_{42}=0.5	0.5078	0.0535	0.0536	0.0492	1.0895
σ_{43}=0.5	0.5093	0.0617	0.0618	0.0430	1.4355

Table 1a(i).5

Simulation Result of Threshold and Correlation Estimation

Symmetric distribution of Z_1 and Z_2 with small length for Y_1 and Y_2

(N = 400)

Parameter	Mean	RMS	$\overline{\text{S.E.}}$	S.E.	Ratio
μ_1 =0.0	-0.0005	0.0490	0.0497	0.0497	0.9996
μ_2 =0.0	0.0072	0.0516	0.0517	0.0492	1.0522
σ_{33}=1.0	1.0052	0.0791	0.0799	0.0711	1.1242
σ_{44}=1.0	0.9834	0.0706	0.0695	0.0695	1.0000
$\alpha_{1,2}$=-0.5	-0.5119	0.0668	0.0666	0.0647	1.0286
$\alpha_{1,3}$=0.5	0.5165	0.0552	0.0534	0.0603	0.8851
$\alpha_{2,2}$=-0.5	-0.4835	0.0615	0.0600	0.0644	0.9308
$\alpha_{2,3}$=0.5	0.4955	0.0686	0.0693	0.0602	1.1515
ρ_{21}=0.5	0.5037	0.0534	0.0540	0.0489	1.1032
ρ_{31}=0.5	0.4890	0.0429	0.0420	0.0435	0.9661
ρ_{32}=0.5	0.5077	0.0407	0.0405	0.0426	0.9495
ρ_{41}=0.5	0.4909	0.0524	0.0524	0.0433	1.2093
ρ_{42}=0.5	0.5018	0.0393	0.0397	0.0431	0.9222
σ_{43}=0.5	0.4926	0.0539	0.0541	0.0374	1.4473

Table 1a(ii).1

Simulation Result of Threshold and Correlation Estimation

Symmetric distribution of Z_1 and Z_2

with intermediate length for Y_1 and Y_2

(N = 70)

Parameter	Mean	RMS	$\overline{\text{S.E.}}$	S.E.	Ratio
μ_1 =0.0	-0.0121	0.1317	0.1329	0.1202	1.1057
μ_2 =0.0	-0.0088	0.1108	0.1119	0.1215	0.9206
σ_{33}=1.0	0.9749	0.1539	0.1537	0.1671	0.9201
σ_{44}=1.0	1.0057	0.1681	0.1702	0.1722	0.9881
$\alpha_{1,2}$=-0.5	-0.4924	0.1887	0.1909	0.1584	1.2051
$\alpha_{1,3}$=0.5	0.5100	0.1332	0.1345	0.1456	0.9234
$\alpha_{2,2}$=-0.5	-0.4801	0.1278	0.1279	0.1568	0.8157
$\alpha_{2,3}$=0.5	0.5281	0.1845	0.1847	0.1464	1.2613
ρ_{21}=0.5	0.4718	0.1246	0.1229	0.1204	1.0216
ρ_{31}=0.5	0.5025	0.1076	0.1090	0.1029	1.0586
ρ_{32}=0.5	0.4809	0.1161	0.1160	0.1056	1.0983
ρ_{41}=0.5	0.4996	0.0947	0.0960	0.1038	0.9245
ρ_{42}=0.5	0.4956	0.0967	0.0979	0.1041	0.9402
σ_{43}=0.5	0.4875	0.1159	0.1166	0.0889	1.3118

Table 1a(ii).2

Simulation Result of Threshold and Correlation Estimation

Symmetric distribution of Z_1 and Z_2

with intermediate length for Y_1 and Y_2

(N = 100)

Parameter	Mean	RMS	$\overline{\text{S.E.}}$	S.E.	Ratio
μ_1 =0.0	-0.0188	0.1076	0.1073	0.1023	1.0492
μ_2 =0.0	0.0023	0.1097	0.1111	0.1015	1.0940
σ_{33}=1.0	1.0408	0.1470	0.1430	0.1490	0.9597
σ_{44}=1.0	1.0099	0.1593	0.1610	0.1447	1.1128
$\alpha_{1,2}$=-0.5	-0.4749	0.1552	0.1551	0.1306	1.1874
$\alpha_{1,3}$=0.5	0.5159	0.1312	0.1319	0.1211	1.0897
$\alpha_{2,2}$=-0.5	-0.5187	0.1312	0.1315	0.1320	0.9963
$\alpha_{2,3}$=0.5	0.5047	0.1217	0.1231	0.1212	1.0158
ρ_{21}=0.5	0.5236	0.1019	0.1004	0.0953	1.0532
ρ_{31}=0.5	0.5191	0.0787	0.0774	0.0852	0.9077
ρ_{32}=0.5	0.5193	0.0625	0.0602	0.0852	0.7060
ρ_{41}=0.5	0.5146	0.0790	0.0787	0.0853	0.9224
ρ_{42}=0.5	0.5064	0.0729	0.0736	0.0861	0.8549
σ_{43}=0.5	0.5434	0.1380	0.1326	0.0727	1.8250

Table 1a(ii).3

Simulation Result of Threshold and Correlation Estimation

Symmetric distribution of Z_1 and Z_2

with intermediate length for Y_1 and Y_2

(N = 200)

Parameter	Mean	RMS	$\overline{\text{S.E.}}$	S.E.	Ratio
μ_1 =0.0	0.0023	0.0617	0.0624	0.0706	0.8842
μ_2 =0.0	-0.0059	0.0685	0.0691	0.0712	0.9704
σ_{33}=1.0	0.9819	0.0926	0.0919	0.0995	0.9243
σ_{44}=1.0	1.0036	0.0804	0.0813	0.1016	0.8000
$\alpha_{1,2}$=-0.5	-0.5012	0.0825	0.0836	0.0926	0.9020
$\alpha_{1,3}$=0.5	0.5131	0.0797	0.0796	0.0854	0.9321
$\alpha_{2,2}$=-0.5	-0.5040	0.0850	0.0860	0.0918	0.9366
$\alpha_{2,3}$=0.5	0.4972	0.0925	0.0937	0.0855	1.0954
ρ_{21}=0.5	0.5001	0.0720	0.0730	0.0693	1.0522
ρ_{31}=0.5	0.5067	0.0751	0.0757	0.0606	1.2494
ρ_{32}=0.5	0.5012	0.0834	0.0844	0.0611	1.3819
ρ_{41}=0.5	0.4946	0.0594	0.0599	0.0615	0.9736
ρ_{42}=0.5	0.4972	0.0594	0.0601	0.0615	0.9772
σ_{43}=0.5	0.4990	0.0709	0.0718	0.0523	1.3745

Table 1a(ii).4

Simulation Result of Threshold and Correlation Estimation

Symmetric distribution of Z_1 and Z_2

with intermediate length for Y_1 and Y_2

(N = 300)

Parameter	Mean	RMS	$\overline{\text{S.E.}}$	S.E.	Ratio
μ_1 =0.0	-0.0044	0.0581	0.0587	0.0574	1.0217
μ_2 =0.0	-0.0037	0.0501	0.0506	0.0576	0.8795
σ_{33}=1.0	0.9897	0.0931	0.0937	0.0818	1.1451
σ_{44}=1.0	0.9953	0.0908	0.0919	0.0823	1.1170
$\alpha_{1,2}$=-0.5	-0.4820	0.0765	0.0753	0.0743	1.0136
$\alpha_{1,3}$=0.5	0.5153	0.0822	0.0818	0.0697	1.1747
$\alpha_{2,2}$=-0.5	-0.5272	0.0696	0.0649	0.0754	0.8604
$\alpha_{2,3}$=0.5	0.4969	0.0878	0.0888	0.0698	1.2753
ρ_{21}=0.5	0.5219	0.0623	0.0591	0.0553	1.0679
ρ_{31}=0.5	0.5062	0.0441	0.0442	0.0497	0.8886
ρ_{32}=0.5	0.5131	0.0470	0.0457	0.0492	0.9292
ρ_{41}=0.5	0.5112	0.0550	0.0545	0.0496	1.0988
ρ_{42}=0.5	0.5024	0.0468	0.0474	0.0500	0.9474
σ_{43}=0.5	0.5109	0.0692	0.0692	0.0420	1.6487

Table 1a(ii).5

Simulation Result of Threshold and Correlation Estimation

Symmetric distribution of Z_1 and Z_2

with intermediate length for Y_1 and Y_2

(N = 400)

Parameter	Mean	RMS	$\overline{\text{S.E.}}$	S.E.	Ratio
μ_1 =0.0	-0.0072	0.0476	0.0476	0.0494	0.9632
μ_2 =0.0	0.0034	0.0441	0.0445	0.0495	0.8984
σ_{33}=1.0	0.9783	0.0649	0.0620	0.0700	0.8848
σ_{44}=1.0	0.9818	0.0742	0.0728	0.0703	1.0357
$\alpha_{1,2}$=-0.5	-0.5096	0.0668	0.0670	0.0647	1.0345
$\alpha_{1,3}$=0.5	0.4947	0.0566	0.0571	0.0602	0.9476
$\alpha_{2,2}$=-0.5	-0.4907	0.0810	0.0815	0.0648	1.2575
$\alpha_{2,3}$=0.5	0.5142	0.0653	0.0646	0.0603	1.0707
ρ_{21}=0.5	0.5064	0.0510	0.0513	0.0488	1.0510
ρ_{31}=0.5	0.5043	0.0423	0.0427	0.0431	0.9895
ρ_{32}=0.5	0.5064	0.0388	0.0388	0.0430	0.9020
ρ_{41}=0.5	0.4965	0.0435	0.0439	0.0435	1.0087
ρ_{42}=0.5	0.4900	0.0388	0.0380	0.0439	0.8643
σ_{43}=0.5	0.4939	0.0528	0.0531	0.0366	1.4515

Table 1b(i).1

Simulation Result of Threshold and Correlation Estimation

Asymmetric distribution of Z_1 and Z_2 with small length for Y_1 and Y_2

(N = 70)

Parameter	Mean	RMS	$\overline{S.E.}$	S.E.	Ratio
μ_1 =0.0	0.0041	0.1172	0.1186	0.1190	0.9967
μ_2 =0.0	0.0003	0.1346	0.1363	0.1187	1.1481
σ_{33}=1.0	0.9891	0.1821	0.1841	0.1673	1.1000
σ_{44}=1.0	0.9885	0.1985	0.2007	0.1672	1.2003
$\alpha_{1,2}$=0.0	0.0038	0.1379	0.1396	0.1482	0.9423
$\alpha_{1,3}$=1.0	1.0254	0.1740	0.1744	0.1605	1.0862
$\alpha_{2,2}$=0.0	-0.0192	0.1328	0.1331	0.1453	0.9164
$\alpha_{2,3}$=0.5	0.4926	0.1325	0.1340	0.1454	0.9215
ρ_{21}=0.5	0.5228	0.1094	0.1083	0.1231	0.8802
ρ_{31}=0.5	0.4865	0.0999	0.1003	0.1072	0.9358
ρ_{32}=0.5	0.4753	0.0994	0.0975	0.1107	0.8812
ρ_{41}=0.5	0.5174	0.0882	0.0876	0.1048	0.8356
ρ_{42}=0.5	0.4846	0.1130	0.1134	0.1105	1.0257
σ_{43}=0.5	0.5003	0.1561	0.1581	0.0858	1.8430

Table 1b(i).2

Simulation Result of Threshold and Correlation Estimation

Asymmetric distribution of Z_1 and Z_2 with small length for Y_1 and Y_2

(N = 100)

Parameter	Mean	RMS	$\overline{\text{S.E.}}$	S.E.	Ratio
μ_1 =0.0	0.0123	0.0828	0.0830	0.1003	0.8275
μ_2 =0.0	0.0069	0.0900	0.0909	0.0977	0.9299
σ_{33}=1.0	1.0105	0.1335	0.1348	0.1430	0.9430
σ_{44}=1.0	0.9693	0.1800	0.1800	0.1371	1.3104
$\alpha_{1,2}$=0.0	-0.0207	0.1285	0.1284	0.1226	1.0476
$\alpha_{1,3}$=1.0	0.9919	0.1548	0.1566	0.1326	1.1815
$\alpha_{2,2}$=0.0	0.0118	0.1149	0.1157	0.1219	0.9493
$\alpha_{2,3}$=0.5	0.4863	0.1371	0.1382	0.1214	1.1381
ρ_{21}=0.5	0.5219	0.0996	0.0984	0.1034	0.9519
ρ_{31}=0.5	0.5237	0.1112	0.1100	0.0852	1.2912
ρ_{32}=0.5	0.5128	0.0969	0.0972	0.0894	1.0867
ρ_{41}=0.5	0.5236	0.0922	0.0903	0.0855	1.0555
ρ_{42}=0.5	0.4842	0.1144	0.1148	0.0913	1.2577
σ_{43}=0.5	0.4954	0.1181	0.1195	0.0729	1.6395

Table 1b(i).3

Simulation Result of Threshold and Correlation Estimation

Asymmetric distribution of Z_1 and Z_2 with small length for Y_1 and Y_2

(N = 200)

Parameter	Mean	RMS	$\overline{\text{S.E.}}$	S.E.	Ratio
$\mu_1=0.0$	0.0073	0.0620	0.0624	0.0699	0.8921
$\mu_2=0.0$	-0.0009	0.0685	0.0691	0.0707	0.9774
$\sigma_{33}=1.0$	0.9898	0.0926	0.0891	0.0989	0.9003
$\sigma_{44}=1.0$	1.0127	0.0804	0.0812	0.1012	0.8016
$\alpha_{1,2}=0.0$	-0.0001	0.0922	0.0934	0.0868	1.0761
$\alpha_{1,3}=1.0$	1.0220	0.1070	0.1060	0.0939	1.1294
$\alpha_{2,2}=0.0$	-0.0133	0.0874	0.0875	0.0859	1.0195
$\alpha_{2,3}=0.5$	0.4971	0.0924	0.0935	0.0854	1.0953
$\rho_{21}=0.5$	0.5002	0.0828	0.0839	0.0747	1.1237
$\rho_{31}=0.5$	0.5240	0.0791	0.0757	0.0607	1.2419
$\rho_{32}=0.5$	0.4995	0.0848	0.0842	0.0640	1.3153
$\rho_{41}=0.5$	0.5080	0.0578	0.0575	0.0620	0.9273
$\rho_{42}=0.5$	0.4926	0.0683	0.0674	0.0645	1.0446
$\sigma_{43}=0.5$	0.4945	0.0709	0.0711	0.0532	1.3366

Table 1b(i).4

Simulation Result of Threshold and Correlation Estimation

Asymmetric distribution of Z_1 and Z_2 with small length for Y_1 and Y_2

(N = 300)

Parameter	Mean	RMS	$\overline{\text{S.E.}}$	S.E.	Ratio
μ_1 =0.0	-0.0018	0.0564	0.0571	0.0572	0.9989
μ_2 =0.0	-0.0013	0.0517	0.0523	0.0574	0.9118
σ_{33}=1.0	0.9968	0.0909	0.0920	0.0813	1.1313
σ_{44}=1.0	1.0061	0.0898	0.0908	0.0821	1.1059
$\alpha_{1,2}$=0.0	0.0146	0.0780	0.0775	0.0707	1.0969
$\alpha_{1,3}$=1.0	1.0154	0.0912	0.0911	0.0764	1.1925
$\alpha_{2,2}$=0.0	0.0003	0.0613	0.0621	0.0701	0.8864
$\alpha_{2,3}$=0.5	0.4946	0.0853	0.0862	0.0696	1.2391
ρ_{21}=0.5	0.5210	0.0721	0.0698	0.0599	1.1655
ρ_{31}=0.5	0.4997	0.0479	0.0485	0.0512	0.9462
ρ_{32}=0.5	0.5042	0.0463	0.0467	0.0521	0.8949
ρ_{41}=0.5	0.4971	0.0563	0.0569	0.0516	1.1033
ρ_{42}=0.5	0.4995	0.0539	0.0546	0.0526	1.0377
σ_{43}=0.5	0.5055	0.0683	0.0690	0.0429	1.6081

Table 1b(i).5

Simulation Result of Threshold and Correlation Estimation

Asymmetric distribution of Z_1 and Z_2 with small length for Y_1 and Y_2

(N = 400)

Parameter	Mean	RMS	$\overline{\text{S.E.}}$	S.E.	Ratio
μ_1 =0.0	-0.0022	0.0471	0.0476	0.0492	0.9677
μ_2 =0.0	0.0084	0.0447	0.0445	0.0493	0.9020
σ_{33}=1.0	0.9871	0.0621	0.0615	0.0698	0.8814
σ_{44}=1.0	0.9893	0.0723	0.0725	0.0699	1.0362
$\alpha_{1,2}$=0.0	-0.0024	0.0574	0.0580	0.0613	0.9465
$\alpha_{1,3}$=1.0	0.9863	0.0682	0.0677	0.0655	1.0342
$\alpha_{2,2}$=0.0	0.0031	0.0660	0.0667	0.0605	1.1037
$\alpha_{2,3}$=0.5	0.5148	0.0649	0.0640	0.0603	1.0609
ρ_{21}=0.5	0.5098	0.0591	0.0590	0.0524	1.1276
ρ_{31}=0.5	0.5025	0.0437	0.0442	0.0441	1.0017
ρ_{32}=0.5	0.5109	0.0476	0.0470	0.0448	1.0491
ρ_{41}=0.5	0.4993	0.0483	0.0489	0.0443	1.1047
ρ_{42}=0.5	0.4901	0.0463	0.0458	0.0460	0.9954
σ_{43}=0.5	0.4890	0.0527	0.0522	0.0372	1.4027

Table 1b(ii).1

Simulation Result of Threshold and Correlation Estimation

Asymmetric distribution of Z_1 and Z_2

with intermediate length for Y_1 and Y_2

(N = 70)

Parameter	Mean	RMS	$\overline{\text{S.E.}}$	S.E.	Ratio
μ_1 =0.0	0.0004	0.1227	0.1243	0.1179	1.0538
μ_2 =0.0	-0.0015	0.1034	0.1047	0.1206	0.8686
σ_{33}=1.0	0.9583	0.1455	0.1412	0.1642	0.8601
σ_{44}=1.0	0.9934	0.1713	0.1733	0.1701	1.0191
$\alpha_{1,2}$=0.0	-0.0003	0.1360	0.1377	0.1482	0.9288
$\alpha_{1,3}$=1.0	1.0153	0.2130	0.2152	0.1611	1.3362
$\alpha_{2,2}$=0.0	0.0186	0.1263	0.1265	0.1461	0.8657
$\alpha_{2,3}$=0.5	0.5126	0.1707	0.1724	0.1461	1.1796
ρ_{21}=0.5	0.4730	0.1505	0.1500	0.1291	1.1614
ρ_{31}=0.5	0.4926	0.0898	0.0906	0.1081	0.8380
ρ_{32}=0.5	0.4930	0.0993	0.1003	0.1104	0.9087
ρ_{41}=0.5	0.4921	0.1043	0.1053	0.1077	0.9779
ρ_{42}=0.5	0.4911	0.0978	0.0986	0.1106	0.8918
σ_{43}=0.5	0.4783	0.1239	0.1236	0.0879	1.4059

Table 1b(ii).2

Simulation Result of Threshold and Correlation Estimation

Asymmetric distribution of Z_1 and Z_2

with intermediate length for Y_1 and Y_2

(N = 100)

Parameter	Mean	RMS	$\overline{\text{S.E.}}$	S.E.	Ratio
μ_1 =0.0	-0.0229	0.1203	0.1196	0.1032	1.1580
μ_2 =0.0	-0.0042	0.1155	0.1168	0.1013	1.1532
σ_{33}=1.0	1.0483	0.1483	0.1420	0.1500	0.9464
σ_{44}=1.0	0.9985	0.1332	0.1349	0.1429	0.9438
$\alpha_{1,2}$=0.0	-0.0030	0.1620	0.1640	0.1253	1.3089
$\alpha_{1,3}$=1.0	1.0261	0.1523	0.1519	0.1340	1.1334
$\alpha_{2,2}$=0.0	0.0200	0.1082	0.1077	0.1217	0.8851
$\alpha_{2,3}$=0.5	0.5123	0.1212	0.1221	0.1216	1.0042
ρ_{21}=0.5	0.4955	0.1179	0.1193	0.1060	1.1252
ρ_{31}=0.5	0.5145	0.0676	0.0669	0.0882	0.7578
ρ_{32}=0.5	0.5132	0.0654	0.0649	0.0910	0.7131
ρ_{41}=0.5	0.5293	0.0789	0.0742	0.0865	0.8572
ρ_{42}=0.5	0.5276	0.0750	0.0706	0.0891	0.7927
σ_{43}=0.5	0.5384	0.1285	0.1242	0.0731	1.7000

Table 1b(ii).3

Simulation Result of Threshold and Correlation Estimation

Asymmetric distribution of Z_1 and Z_2

with intermediate length for Y_1 and Y_2

(N = 200)

Parameter	Mean	RMS	$\overline{\text{S.E.}}$	S.E.	Ratio
μ_1 =0.0	0.0023	0.0617	0.0624	0.0701	0.8900
μ_2 =0.0	-0.0059	0.0685	0.0691	0.0708	0.9754
σ_{33}=1.0	0.9819	0.0926	0.0919	0.0994	0.9250
σ_{44}=1.0	1.0036	0.0804	0.0813	0.1016	0.8005
$\alpha_{1,2}$=0.0	-0.0002	0.0922	0.0934	0.0868	1.0759
$\alpha_{1,3}$=1.0	1.0220	0.1070	0.1060	0.0939	1.1287
$\alpha_{2,2}$=0.0	-0.0133	0.0874	0.0875	0.0858	1.0190
$\alpha_{2,3}$=0.5	0.4971	0.0924	0.0936	0.0854	1.0956
ρ_{21}=0.5	0.5002	0.0828	0.0839	0.0747	1.1237
ρ_{31}=0.5	0.5259	0.0791	0.0757	0.0611	1.2389
ρ_{32}=0.5	0.5010	0.0848	0.0859	0.0645	1.3320
ρ_{41}=0.5	0.5102	0.0578	0.0576	0.0624	0.9233
ρ_{42}=0.5	0.4947	0.0683	0.0690	0.0650	1.0618
σ_{43}=0.5	0.4990	0.0709	0.0718	0.0523	1.3745

Table 1b(ii).4

Simulation Result of Threshold and Correlation Estimation

Asymmetric distribution of Z_1 and Z_2

with intermediate length for Y_1 and Y_2

(N = 300)

Parameter	Mean	RMS	$\overline{\text{S.E.}}$	S.E.	Ratio
μ_1 =0.0	-0.0068	0.0568	0.0571	0.0573	0.9967
μ_2 =0.0	-0.0063	0.0521	0.0523	0.0576	0.9091
σ_{33}=1.0	0.9881	0.0927	0.0931	0.0816	1.1412
σ_{44}=1.0	0.9994	0.0909	0.0921	0.0825	1.1155
$\alpha_{1,2}$=0.0	0.0147	0.0779	0.0775	0.0707	1.0964
$\alpha_{1,3}$=1.0	1.0154	0.0912	0.0910	0.0764	1.1922
$\alpha_{2,2}$=0.0	0.0003	0.0613	0.0621	0.0701	0.8861
$\alpha_{2,3}$=0.5	0.4946	0.0853	0.0862	0.0696	1.2397
ρ_{21}=0.5	0.5210	0.0721	0.0698	0.0599	1.1655
ρ_{31}=0.5	0.5021	0.0468	0.0474	0.0516	0.9186
ρ_{32}=0.5	0.5068	0.0468	0.0469	0.0525	0.8932
ρ_{41}=0.5	0.4993	0.0547	0.0554	0.0520	1.0660
ρ_{42}=0.5	0.5019	0.0547	0.0554	0.0530	1.0458
σ_{43}=0.5	0.5110	0.0692	0.0692	0.0421	1.6454

Table 1b(ii).5

Simulation Result of Threshold and Correlation Estimation

Asymmetric distribution of Z_1 and Z_2

with intermediate length for Y_1 and Y_2

(N = 400)

Parameter	Mean	RMS	$\overline{\text{S.E.}}$	S.E.	Ratio
μ_1 =0.0	-0.0072	0.0476	0.0475	0.0493	0.9656
μ_2 =0.0	0.0034	0.0441	0.0445	0.0495	0.8995
σ_{33}=1.0	0.9783	0.0649	0.0620	0.0700	0.8852
σ_{44}=1.0	0.9818	0.0742	0.0728	0.0703	1.0364
$\alpha_{1,2}$=0.0	-0.0024	0.0573	0.0580	0.0613	0.9459
$\alpha_{1,3}$=1.0	0.9864	0.0682	0.0677	0.0655	1.0342
$\alpha_{2,2}$=0.0	0.0031	0.0660	0.0668	0.0605	1.1040
$\alpha_{2,3}$=0.5	0.5148	0.0649	0.0640	0.0603	1.0616
ρ_{21}=0.5	0.5098	0.0591	0.0590	0.0524	1.1276
ρ_{31}=0.5	0.5052	0.0436	0.0439	0.0444	0.9893
ρ_{32}=0.5	0.5123	0.0485	0.0475	0.0451	1.0533
ρ_{41}=0.5	0.5016	0.0486	0.0492	0.0445	1.1052
ρ_{42}=0.5	0.4909	0.0465	0.0462	0.0463	0.9982
σ_{43}=0.5	0.4939	0.0528	0.0531	0.0366	1.4515

Table 2(i)

Covariance Structure Analysis (Pseudo Estimation)

(N = 300)

Parameter	Mean	RMS	S.E.	$\overline{\text{S.E.}}$	Ratio	MaxD
F(1,1)=0.8	0.8111	0.0453	0.0344	0.0445	1.2926	0.0199
F(2,1)=0.8	0.8118	0.0410	0.0347	0.0398	1.1463	0.0113
F(3,1)=0.8	0.8161	0.0368	0.0343	0.0335	0.9762	0.0077
F(4,1)=0.8	0.8021	0.0332	0.0355	0.0336	0.9462	0.0112
F(5,2)=0.8	0.8154	0.0314	0.0319	0.0278	0.8708	0.0111
F(6,2)=0.8	0.8111	0.0390	0.0321	0.0379	1.1785	0.0125
F(7,2)=0.8	0.8079	0.0528	0.0268	0.0529	1.9755	0.0322
F(8,2)=0.8	0.8056	0.0485	0.0269	0.0488	1.8166	0.0300
M(2,1)=0.5	0.5268	0.0619	0.0526	0.0565	1.0740	0.0125
E(1,1)=0.36	0.3743	0.0885	0.2175	0.0885	0.4069	0.1546
E(2,2)=0.36	0.3768	0.0787	0.2197	0.0779	0.3546	0.1711
E(3,3)=0.36	0.3784	0.0613	0.2199	0.0592	0.2693	0.1844
E(4,4)=0.36	0.3974	0.0807	0.2162	0.0724	0.3349	0.1776
E(5,5)=0.36	0.3598	0.0630	0.2184	0.0638	0.2919	0.1799
E(6,6)=0.36	0.3632	0.0699	0.2182	0.0708	0.3242	0.1638
E(7,7)=0.36	0.3620	0.0408	0.0718	0.0413	0.5749	0.0434
E(8,8)=0.36	0.3730	0.0544	0.0717	0.0535	0.7466	0.0336

P-value for H_o = 0.9917

Table 2(ii)

Covariance Structure Analysis (Pseudo Estimation)

(N = 400)

Parameter	Mean	RMS	S.E.	$\overline{\text{S.E.}}$	Ratio	MaxD
F(1,1)=0.8	0.8074	0.0297	0.0301	0.0291	0.9652	0.0064
F(2,1)=0.8	0.8168	0.0328	0.0296	0.0285	0.9634	0.0087
F(3,1)=0.8	0.8209	0.0347	0.0293	0.0281	0.9599	0.0071
F(4,1)=0.8	0.8116	0.0368	0.0299	0.0354	1.1828	0.0120
F(5,2)=0.8	0.8083	0.0323	0.0282	0.0316	1.1201	0.0105
F(6,2)=0.8	0.8071	0.0242	0.0282	0.0235	0.8317	0.0102
F(7,2)=0.8	0.8084	0.0430	0.0231	0.0427	1.8466	0.0247
F(8,2)=0.8	0.8090	0.0493	0.0233	0.0491	2.1048	0.0291
M(2,1)=0.5	0.5205	0.0593	0.0459	0.0564	1.2290	0.0158
E(1,1)=0.36	0.3735	0.0609	0.1890	0.0601	0.3181	0.1466
E(2,2)=0.36	0.3587	0.0524	0.1927	0.0531	0.2756	0.1559
E(3,3)=0.36	0.3518	0.0517	0.1909	0.0517	0.2706	0.1518
E(4,4)=0.36	0.3665	0.0684	0.1900	0.0690	0.3632	0.1346
E(5,5)=0.36	0.3601	0.0562	0.1908	0.0569	0.2983	0.1546
E(6,6)=0.36	0.3683	0.0518	0.1898	0.0517	0.2725	0.1526
E(7,7)=0.36	0.3561	0.0372	0.0630	0.0375	0.5946	0.0355
E(8,8)=0.36	0.3594	0.0349	0.0626	0.0353	0.5636	0.0404

P-value for H_o = 0.2095

Table 2(iii)

Covariance Structure Analysis (Pseudo Estimation)

(N = 500)

Parameter	Mean	RMS	S.E.	$\overline{\text{S.E.}}$	Ratio	MaxD
F(1,1)=0.8	0.8002	0.0334	0.0272	0.0339	1.2457	0.0126
F(2,1)=0.8	0.8110	0.0320	0.0270	0.0304	1.1355	0.0080
F(3,1)=0.8	0.8159	0.0281	0.0270	0.0234	0.8701	0.0080
F(4,1)=0.8	0.8146	0.0310	0.0271	0.0277	1.0277	0.0052
F(5,2)=0.8	0.8017	0.0224	0.0253	0.0227	0.8944	0.0082
F(6,2)=0.8	0.8102	0.0270	0.0250	0.0253	1.0133	0.0044
F(7,2)=0.8	0.7929	0.0421	0.0213	0.0420	1.9686	0.0241
F(8,2)=0.8	0.8045	0.0370	0.0214	0.0372	1.7355	0.0205
M(2,1)=0.5	0.5107	0.0462	0.0422	0.0456	1.0799	0.0074
E(1,1)=0.36	0.3658	0.0594	0.1706	0.0599	0.3511	0.1294
E(2,2)=0.36	0.3598	0.0456	0.1704	0.0462	0.2713	0.1404
E(3,3)=0.36	0.3396	0.0503	0.1723	0.0465	0.2698	0.1364
E(4,4)=0.36	0.3495	0.0493	0.1701	0.0488	0.2869	0.1349
E(5,5)=0.36	0.3726	0.0508	0.1680	0.0498	0.2964	0.1304
E(6,6)=0.36	0.3552	0.0483	0.1708	0.0487	0.2840	0.1413
E(7,7)=0.36	0.3672	0.0410	0.0557	0.0408	0.7332	0.0229
E(8,8)=0.36	0.3560	0.0378	0.0569	0.0381	0.6689	0.0276

P-value for H_o = 0.9558

Table 2(iv)

Covariance Structure Analysis (Pseudo Estimation)

(N = 700)

Parameter	Mean	RMS	S.E.	$\overline{\text{S.E.}}$	Ratio	MaxD
F(1,1)=0.8	0.8028	0.0263	0.0234	0.0265	1.1323	0.0082
F(2,1)=0.8	0.8078	0.0231	0.0234	0.0220	0.9398	0.0051
F(3,1)=0.8	0.8051	0.0274	0.0236	0.0273	1.1572	0.0086
F(4,1)=0.8	0.8074	0.0253	0.0235	0.0245	1.0404	0.0050
F(5,2)=0.8	0.8032	0.0231	0.0219	0.0232	1.0604	0.0042
F(6,2)=0.8	0.7992	0.0243	0.0223	0.0246	1.1046	0.0056
F(7,2)=0.8	0.8126	0.0328	0.0180	0.0307	1.7025	0.0149
F(8,2)=0.8	0.8013	0.0255	0.0179	0.0258	1.4411	0.0107
M(2,1)=0.5	0.4948	0.0398	0.0366	0.0400	1.0917	0.0071
E(1,1)=0.36	0.3688	0.0487	0.1454	0.0485	0.3339	0.1065
E(2,2)=0.36	0.3584	0.0395	0.1459	0.0400	0.2743	0.1133
E(3,3)=0.36	0.3658	0.0439	0.1449	0.0441	0.3041	0.1108
E(4,4)=0.36	0.3586	0.0474	0.1457	0.0480	0.3292	0.1120
E(5,5)=0.36	0.3617	0.0385	0.1442	0.0390	0.2702	0.1129
E(6,6)=0.36	0.3678	0.0409	0.1439	0.0406	0.2822	0.1149
E(7,7)=0.36	0.3585	0.0309	0.0488	0.0313	0.6402	0.0250
E(8,8)=0.36	0.3595	0.0241	0.0473	0.0244	0.5158	0.0271

P-value for H_o = 0.4739

Table 2(v)

Covariance Structure Analysis (Pseudo Estimation)

(N = 1000)

Parameter	Mean	RMS	S.E.	$\overline{\text{S.E.}}$	Ratio	MaxD
F(1,1)=0.8	0.8015	0.0179	0.0198	0.0181	0.9143	0.0040
F(2,1)=0.8	0.8049	0.0184	0.0196	0.0180	0.9190	0.0040
F(3,1)=0.8	0.8041	0.0188	0.0197	0.0186	0.9407	0.0035
F(4,1)=0.8	0.8065	0.0208	0.0196	0.0200	1.0236	0.0028
F(5,2)=0.8	0.8047	0.0229	0.0183	0.0227	1.2398	0.0068
F(6,2)=0.8	0.8085	0.0214	0.0180	0.0199	1.1052	0.0046
F(7,2)=0.8	0.8094	0.0275	0.0150	0.0262	1.7410	0.0130
F(8,2)=0.8	0.8135	0.0315	0.0149	0.0288	1.9385	0.0159
M(2,1)=0.5	0.5148	0.0348	0.0300	0.0318	1.0600	0.0047
E(1,1)=0.36	0.3649	0.0300	0.1214	0.0300	0.2467	0.0963
E(2,2)=0.36	0.3598	0.0287	0.1214	0.0291	0.2392	0.1004
E(3,3)=0.36	0.3618	0.0307	0.1218	0.0311	0.2550	0.0956
E(4,4)=0.36	0.3568	0.0343	0.1220	0.0346	0.2837	0.0964
E(5,5)=0.36	0.3587	0.0381	0.1210	0.0386	0.3188	0.0880
E(6,6)=0.36	0.3523	0.0342	0.1220	0.0338	0.2767	0.0966
E(7,7)=0.36	0.3597	0.0260	0.0406	0.0263	0.6478	0.0207
E(8,8)=0.36	0.3572	0.0257	0.0412	0.0258	0.6270	0.0207

P-value for H_o = 0.9793

Table 3(i)

Simulation Study on Correlation Structure Analysis

(N = 300)

Parameter	Mean	RMS	S.E.	$\overline{\text{S.E.}}$	Ratio	MaxD
F(1,1)=0.8	0.8122	0.0395	0.0361	0.0381	1.0563	0.0080
F(2,1)=0.8	0.8133	0.0423	0.0363	0.0427	1.1225	0.0152
F(3,1)=0.8	0.8176	0.0433	0.0362	0.0401	1.1078	0.0110
F(4,1)=0.8	0.8151	0.0416	0.0364	0.0393	1.0803	0.0105
F(5,2)=0.8	0.8142	0.0427	0.0362	0.0408	1.1272	0.0129
F(6,2)=0.8	0.8176	0.0437	0.0351	0.0405	1.1547	0.0125
F(7,2)=0.8	0.8076	0.0328	0.0305	0.0323	1.0591	0.0069
F(8,2)=0.8	0.8009	0.0335	0.0306	0.0339	1.1079	0.0092
M(2,1)=0.5	0.5310	0.0751	0.0525	0.0693	1.3201	0.0274
E(1,1)=0.36	0.3390	0.0635	------	------	------	------
E(2,2)=0.36	0.3370	0.0687	------	------	------	------
E(3,3)=0.36	0.3300	0.0711	------	------	------	------
E(4,4)=0.36	0.3341	0.0680	------	------	------	------
E(5,5)=0.36	0.3355	0.0696	------	------	------	------
E(6,6)=0.36	0.3300	0.0723	------	------	------	------
E(7,7)=0.36	0.3467	0.0530	------	------	------	------
E(8,8)=0.36	0.3574	0.0540	------	------	------	------

P-value for H_o = 0.0060

Table 3(ii)

Simulation Study on Correlation Structure Analysis

(N = 400)

Parameter	Mean	RMS	S.E.	$\overline{\text{S.E.}}$	Ratio	MaxD
F(1,1)=0.8	0.8075	0.0308	0.0319	0.0302	0.9458	0.0073
F(2,1)=0.8	0.8164	0.0330	0.0314	0.0390	0.9228	0.0101
F(3,1)=0.8	0.8212	0.0348	0.0311	0.0279	0.8988	0.0092
F(4,1)=0.8	0.8121	0.0328	0.0317	0.0307	1.1581	0.0115
F(5,2)=0.8	0.8079	0.0335	0.0317	0.0330	1.0398	0.0112
F(6,2)=0.8	0.8068	0.0245	0.0315	0.0239	0.7575	0.0130
F(7,2)=0.8	0.8094	0.0252	0.0271	0.0237	0.8766	0.0129
F(8,2)=0.8	0.8079	0.0265	0.0273	0.0257	0.9395	0.0066
M(2,1)=0.5	0.5157	0.0751	0.0467	0.0556	1.1898	0.0140
E(1,1)=0.36	0.3470	0.0498	------	------	------	------
E(2,2)=0.36	0.3326	0.0543	------	------	------	------
E(3,3)=0.36	0.3250	0.0571	------	------	------	------
E(4,4)=0.36	0.3391	0.0619	------	------	------	------
E(5,5)=0.36	0.3462	0.0533	------	------	------	------
E(6,6)=0.36	0.3486	0.0397	------	------	------	------
E(7,7)=0.36	0.3443	0.0410	------	------	------	------
E(8,8)=0.36	0.3466	0.0431	------	------	------	------

P-value for H_o = 0.1017

Table 3(iii)

Simulation Study on Correlation Structure Analysis

(N = 500)

Parameter	Mean	RMS	S.E.	$\overline{\text{S.E.}}$	Ratio	MaxD
F(1,1)=0.8	0.8016	0.0335	0.0290	0.0339	1.1689	0.0106
F(2,1)=0.8	0.8112	0.0325	0.0286	0.0309	1.0821	0.0061
F(3,1)=0.8	0.8147	0.0269	0.0290	0.0228	0.7875	0.0106
F(4,1)=0.8	0.8162	0.0319	0.0287	0.0278	0.9674	0.0058
F(5,2)=0.8	0.8014	0.0237	0.0284	0.0239	0.8418	0.0098
F(6,2)=0.8	0.8100	0.0272	0.0282	0.0256	0.9089	0.0070
F(7,2)=0.8	0.7982	0.0267	0.0251	0.0270	1.0753	0.0059
F(8,2)=0.8	0.8060	0.0250	0.0252	0.0245	0.9756	0.0053
M(2,1)=0.5	0.5085	0.0453	0.0430	0.0451	1.0496	0.0064
E(1,1)=0.36	0.3563	0.0536	------	------	------	------
E(2,2)=0.36	0.3410	0.0525	------	------	------	------
E(3,3)=0.36	0.3357	0.0441	------	------	------	------
E(4,4)=0.36	0.3330	0.0518	------	------	------	------
E(5,5)=0.36	0.3572	0.0379	------	------	------	------
E(6,6)=0.36	0.3433	0.0442	------	------	------	------
E(7,7)=0.36	0.3622	0.0424	------	------	------	------
E(8,8)=0.36	0.3499	0.0405	------	------	------	------

P-value for H_o = 0.0759

Table 3(iv)

Simulation Study on Correlation Structure Analysis

(N = 700)

Parameter	Mean	RMS	S.E.	$\overline{\text{S.E.}}$	Ratio	MaxD
F(1,1)=0.8	0.8055	0.0196	0.0247	0.0190	0.7710	0.0086
F(2,1)=0.8	0.8110	0.0292	0.0250	0.0274	1.0970	0.0060
F(3,1)=0.8	0.8074	0.0315	0.0247	0.0310	1.2522	0.0115
F(4,1)=0.8	0.8117	0.0247	0.0244	0.0220	0.9022	0.0076
F(5,2)=0.8	0.8072	0.0273	0.0241	0.0266	1.1063	0.0064
F(6,2)=0.8	0.8099	0.0258	0.0246	0.0241	0.9836	0.0041
F(7,2)=0.8	0.8094	0.0189	0.0207	0.0166	0.8000	0.0072
F(8,2)=0.8	0.7996	0.0181	0.0213	0.0183	0.8598	0.0064
M(2,1)=0.5	0.5054	0.0355	0.0367	0.0356	0.9706	0.0038
E(1,1)=0.36	0.3508	0.0315	------	------	------	------
E(2,2)=0.36	0.3416	0.0477	------	------	------	------
E(3,3)=0.36	0.3471	0.0509	------	------	------	------
E(4,4)=0.36	0.3407	0.0400	------	------	------	------
E(5,5)=0.36	0.3478	0.0439	------	------	------	------
E(6,6)=0.36	0.3436	0.0420	------	------	------	------
E(7,7)=0.36	0.3447	0.0306	------	------	------	------
E(8,8)=0.36	0.3603	0.0288	------	------	------	------

P-value for H_o = 0.0912

Table 3(v)

Simulation Study on Correlation Structure Analysis

(N = 1000)

Parameter	Mean	RMS	S.E.	$\overline{\text{S.E.}}$	Ratio	MaxD
F(1,1)=0.8	0.8062	0.0172	0.0211	0.0162	0.7702	0.0071
F(2,1)=0.8	0.7999	0.0165	0.0213	0.0167	0.7816	0.0068
F(3,1)=0.8	0.8045	0.0238	0.0209	0.0237	1.1332	0.0055
F(4,1)=0.8	0.8063	0.0204	0.0212	0.0197	0.9302	0.0045
F(5,2)=0.8	0.8003	0.0156	0.0209	0.0158	0.7539	0.0081
F(6,2)=0.8	0.8061	0.0195	0.0207	0.0187	0.9020	0.0050
F(7,2)=0.8	0.7976	0.0136	0.0178	0.0135	0.7575	0.0067
F(8,2)=0.8	0.8002	0.0184	0.0180	0.0187	1.0370	0.0023
M(2,1)=0.5	0.5008	0.0303	0.0311	0.0307	0.9865	0.0047
E(1,1)=0.36	0.3499	0.0277	------	------	------	------
E(2,2)=0.36	0.3598	0.0262	------	------	------	------
E(3,3)=0.36	0.3522	0.0383	------	------	------	------
E(4,4)=0.36	0.3496	0.0329	------	------	------	------
E(5,5)=0.36	0.3593	0.0248	------	------	------	------
E(6,6)=0.36	0.3499	0.0314	------	------	------	------
E(7,7)=0.36	0.3636	0.0215	------	------	------	------
E(8,8)=0.36	0.3594	0.0294	------	------	------	------

P-value for H_o = 0.9528

Table 4.1

Comparison of Two Methods

(N = 100)

Note: the standard errors are given in parenthesis.

Parameters	Three-stage Procedure PML Estimate	Two-stage Procedure PML Estimate
F(1,1)=0.8	0.7189 (0.0601)	0.7229 (0.0618)
F(2,1)=0.8	0.8727 (0.0512)	0.8735 (0.0553)
F(3,1)=0.8	0.8344 (0.0551)	0.8368 (0.0567)
F(4,1)=0.8	0.7806 (0.0703)	0.7754 (0.0702)
F(5,2)=0.8	0.8267 (0.0570)	0.8192 (0.0637)
F(6,2)=0.8	0.7982 (0.0547)	0.8074 (0.0614)
F(7,2)=0.8	0.8685 (0.0334)	0.8197 (0.0352)
F(8,2)=0.8	0.8386 (0.0372)	0.8011 (0.0384)
M(2,1)=0.5	0.5049 (0.0814)	0.4980 (0.0867)
E(1,1)=0.36	0.5094 (0.3572)	0.4774 (------)
E(2,2)=0.36	0.2059 (0.3804)	0.2369 (------)
E(3,3)=0.36	0.1740 (0.4023)	0.2998 (------)
E(4,4)=0.36	0.3344 (0.3679)	0.3987 (------)
E(5,5)=0.36	0.4668 (0.3648)	0.3290 (------)
E(6,6)=0.36	0.3340 (0.3338)	0.3482 (------)
E(7,7)=0.36	0.4153 (0.1174)	0.3281 (------)
E(8,8)=0.36	0.4195 (0.1059)	0.3582 (------)
χ^2 values	6.9897	6.4727

Table 4.2

Comparison of Two Methods

(N = 400)

Note: the standard errors are given in parenthesis.

	Three-stage Procedure	Two-stage Procedure
Parameters	PML Estimate	PML Estimate
F(1,1)=0.8	0.7816 (0.0301)	0.7840 (0.0319)
F(2,1)=0.8	0.8481 (0.0261)	0.8475 (0.0274)
F(3,1)=0.8	0.8433 (0.0288)	0.8389 (0.0309)
F(4,1)=0.8	0.8261 (0.0288)	0.8265 (0.0309)
F(5,2)=0.8	0.7567 (0.0331)	0.7545 (0.0362)
F(6,2)=0.8	0.7704 (0.0300)	0.7772 (0.0322)
F(7,2)=0.8	0.7884 (0.0215)	0.8190 (0.0271)
F(8,2)=0.8	0.8252 (0.0258)	0.8035 (0.0286)
M(2,1)=0.5	0.5791 (0.0405)	0.5765 (0.0416)
E(1,1)=0.36	0.4106 (0.1962)	0.3853 (------)
E(2,2)=0.36	0.3308 (0.2022)	0.2817 (------)
E(3,3)=0.36	0.3360 (0.2021)	0.2963 (------)
E(4,4)=0.36	0.3538 (0.2036)	0.3168 (------)
E(5,5)=0.36	0.4349 (0.3648)	0.4308 (------)
E(6,6)=0.36	0.4287 (0.3338)	0.3960 (------)
E(7,7)=0.36	0.3261 (0.0576)	0.3292 (------)
E(8,8)=0.36	0.3751 (0.0724)	0.3544 (------)
χ^2 values	19.0183	16.2134

Table 4.3

Comparison of Two Methods

(N = 500)

Note: the standard errors are given in parenthesis.

Parameters	Three-stage Procedure PML Estimate	Two-stage Procedure PML Estimate
F(1,1)=0.8	0.8185 (0.0267)	0.8236 (0.0284)
F(2,1)=0.8	0.8184 (0.0260)	0.8170 (0.0276)
F(3,1)=0.8	0.8155 (0.0268)	0.8164 (0.0287)
F(4,1)=0.8	0.7614 (0.0314)	0.7581 (0.0330)
F(5,2)=0.8	0.8627 (0.0210)	0.8625 (0.0236)
F(6,2)=0.8	0.8544 (0.0243)	0.8542 (0.0272)
F(7,2)=0.8	0.8824 (0.0189)	0.8357 (0.0211)
F(8,2)=0.8	0.7932 (0.0217)	0.7853 (0.0240)
M(2,1)=0.5	0.4973 (0.0447)	0.4939 (0.0460)
E(1,1)=0.36	0.3076 (0.1715)	0.3217 (------)
E(2,2)=0.36	0.3419 (0.1660)	0.3324 (------)
E(3,3)=0.36	0.3703 (0.1745)	0.3334 (------)
E(4,4)=0.36	0.4189 (0.1727)	0.4253 (------)
E(5,5)=0.36	0.2415 (0.1707)	0.2560 (------)
E(6,6)=0.36	0.2844 (0.1749)	0.2704 (------)
E(7,7)=0.36	0.3363 (0.0633)	0.3016 (------)
E(8,8)=0.36	0.3775 (0.0561)	0.3834 (------)
χ^2 values	17.2195	15.5726

Table 4.4

Comparison of Two Methods

(N = 1000)

Note: the standard errors are given in parenthesis.

	Three-stage Procedure	Two-stage Procedure
Parameters	PML Estimate	PML Estimate
F(1,1)=0.8	0.7971 (0.0210)	0.7954 (0.0227)
F(2,1)=0.8	0.7900 (0.0202)	0.7905 (0.0215)
F(3,1)=0.8	0.7939 (0.0202)	0.7940 (0.0217)
F(4,1)=0.8	0.8113 (0.0199)	0.8123 (0.0214)
F(5,2)=0.8	0.7609 (0.0218)	0.7612 (0.0238)
F(6,2)=0.8	0.8218 (0.0181)	0.8266 (0.0205)
F(7,2)=0.8	0.7637 (0.0169)	0.7779 (0.0202)
F(8,2)=0.8	0.7881 (0.0151)	0.7856 (0.0172)
M(2,1)=0.5	0.5116 (0.0308)	0.5117 (0.0314)
E(1,1)=0.36	0.3648 (0.1247)	0.3674 (------)
E(2,2)=0.36	0.3785 (0.1181)	0.3753 (------)
E(3,3)=0.36	0.3631 (0.1244)	0.3696 (------)
E(4,4)=0.36	0.3378 (0.1207)	0.3402 (------)
E(5,5)=0.36	0.4328 (0.1188)	0.4205 (------)
E(6,6)=0.36	0.3002 (0.1212)	0.3168 (------)
E(7,7)=0.36	0.3877 (0.0407)	0.3949 (------)
E(8,8)=0.36	0.3782 (0.0391)	0.3828 (------)
χ^2 values	23.8501	21.8269

Figure 1

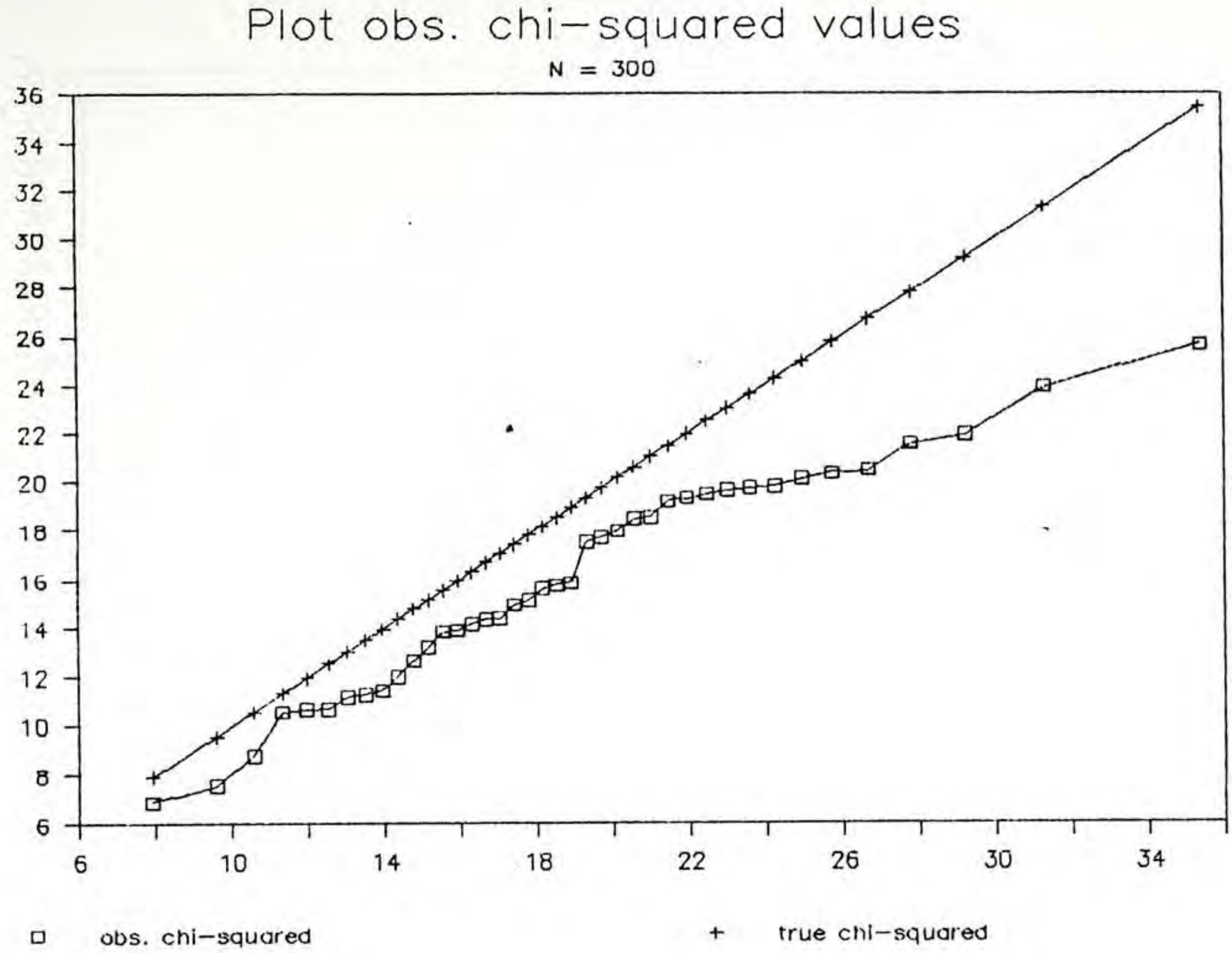


Figure 2

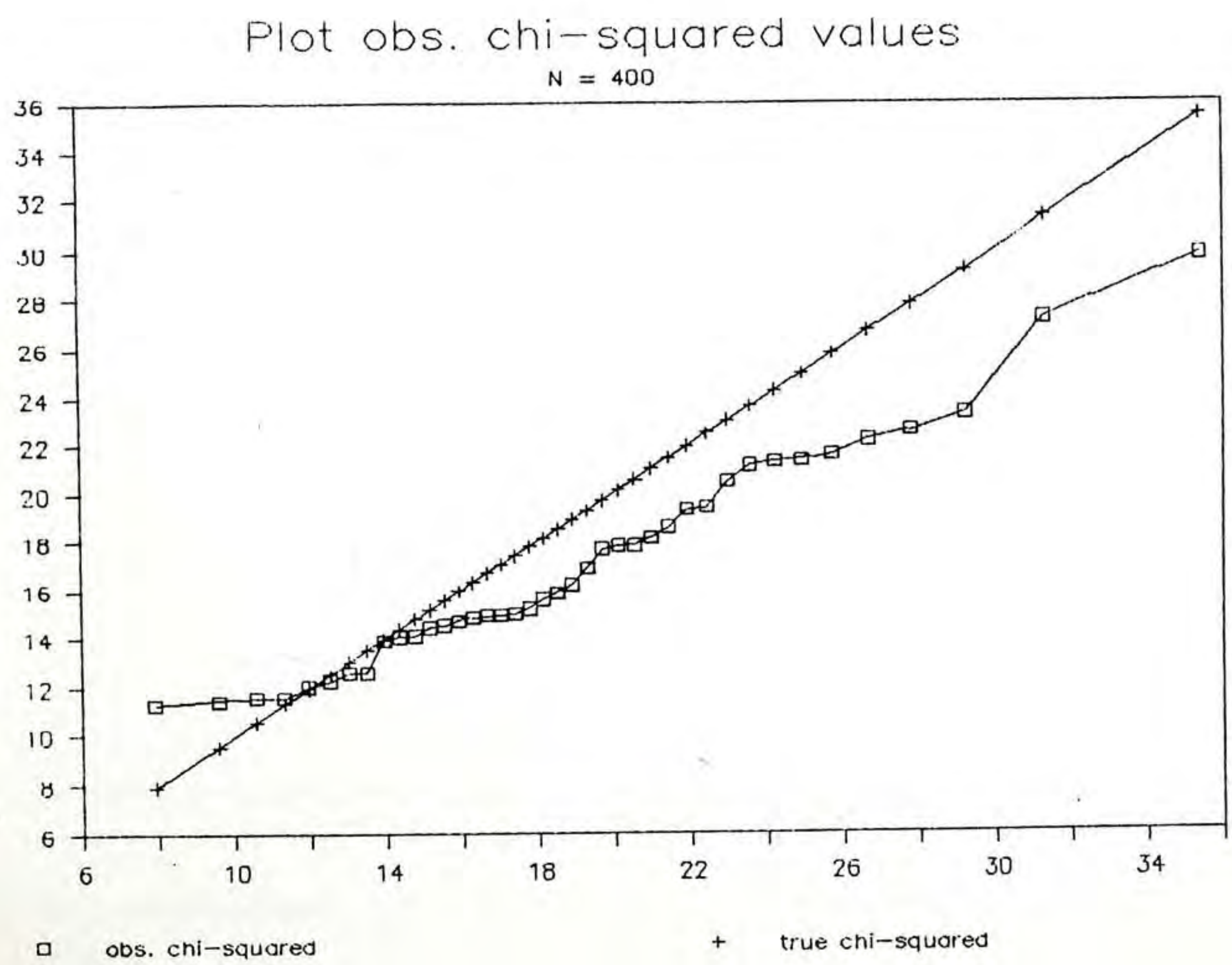


Figure 3
Plot obs. chi-squared values

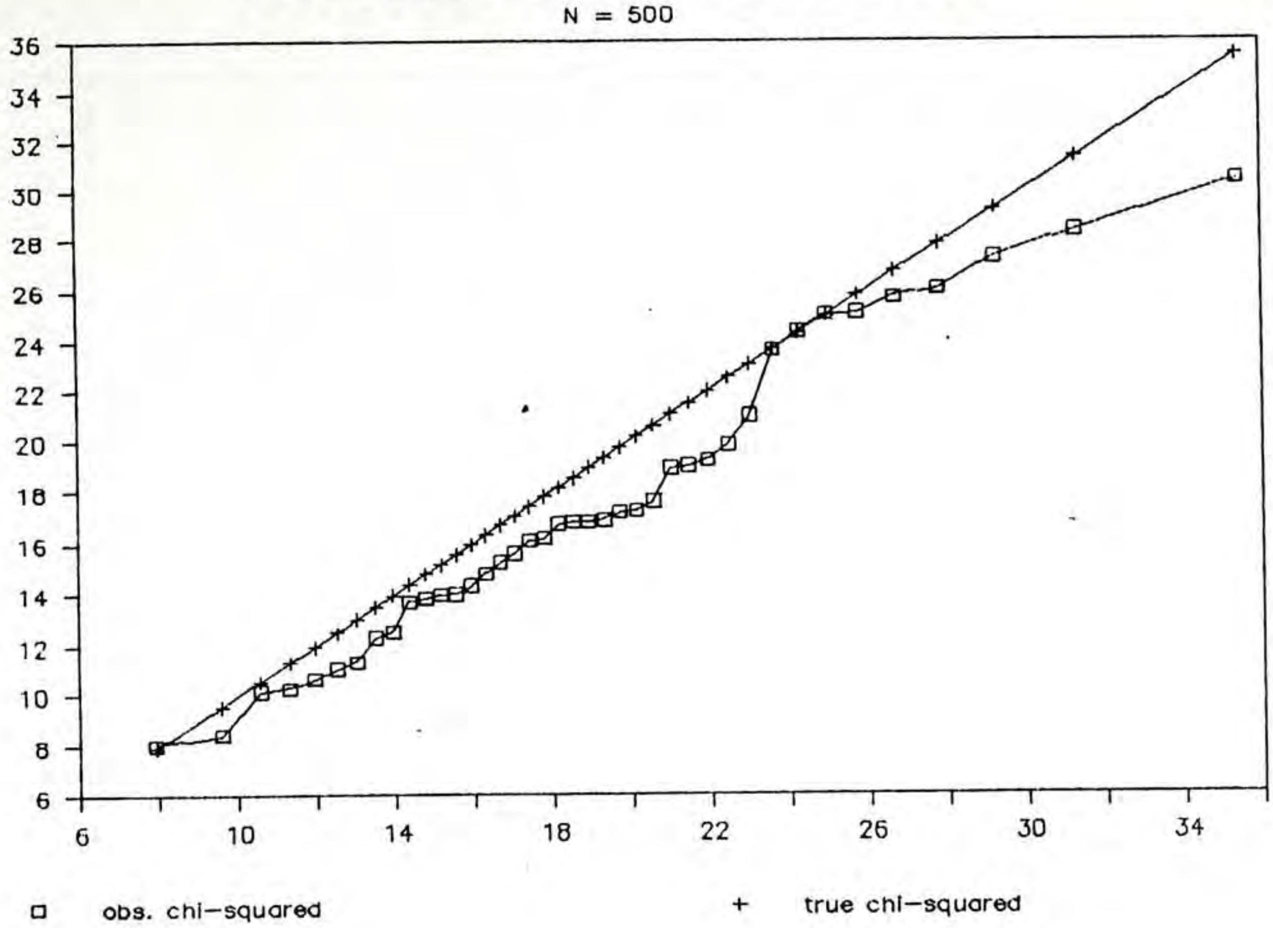


Figure 4
Plot obs. chi-squared values

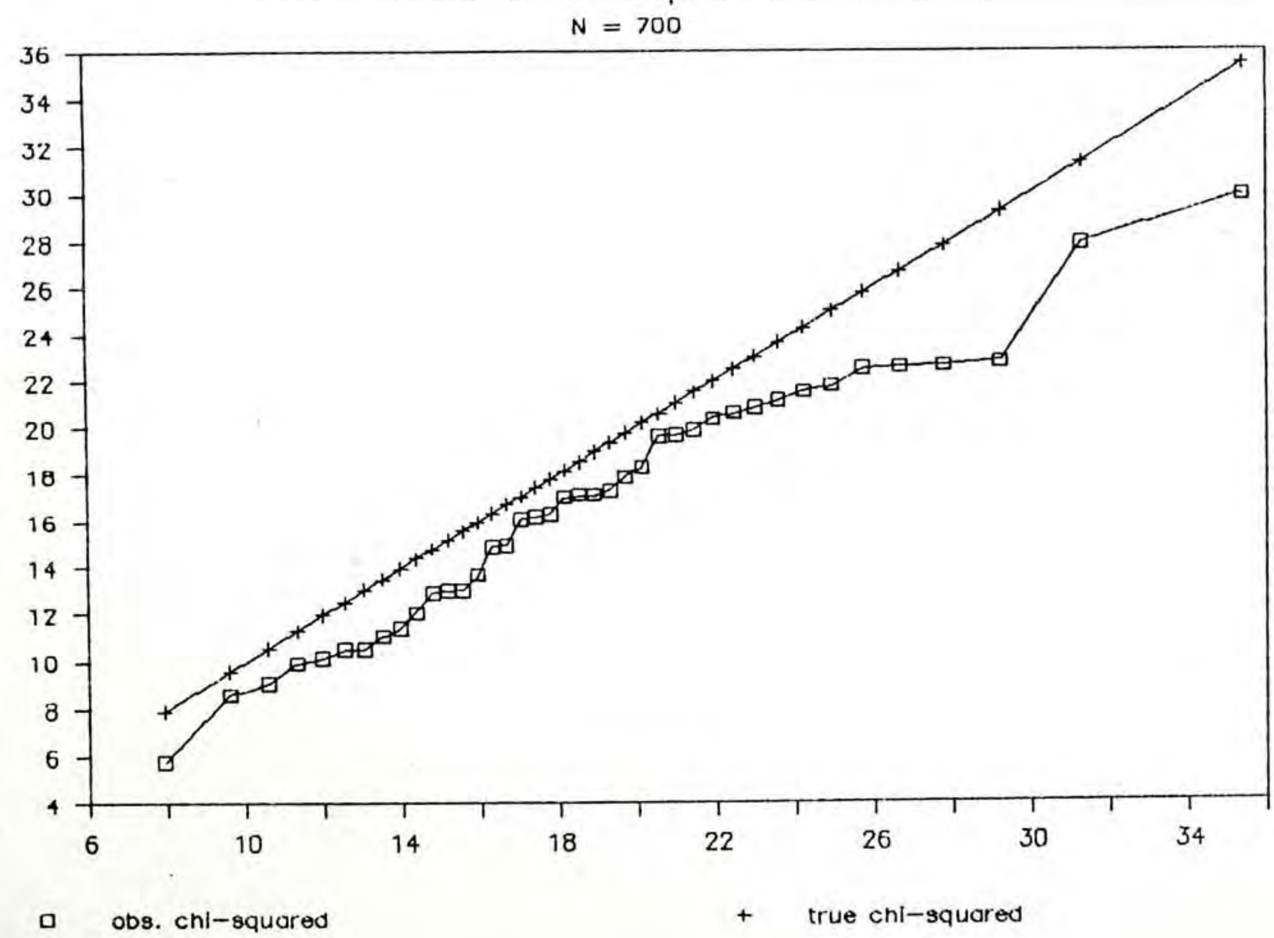
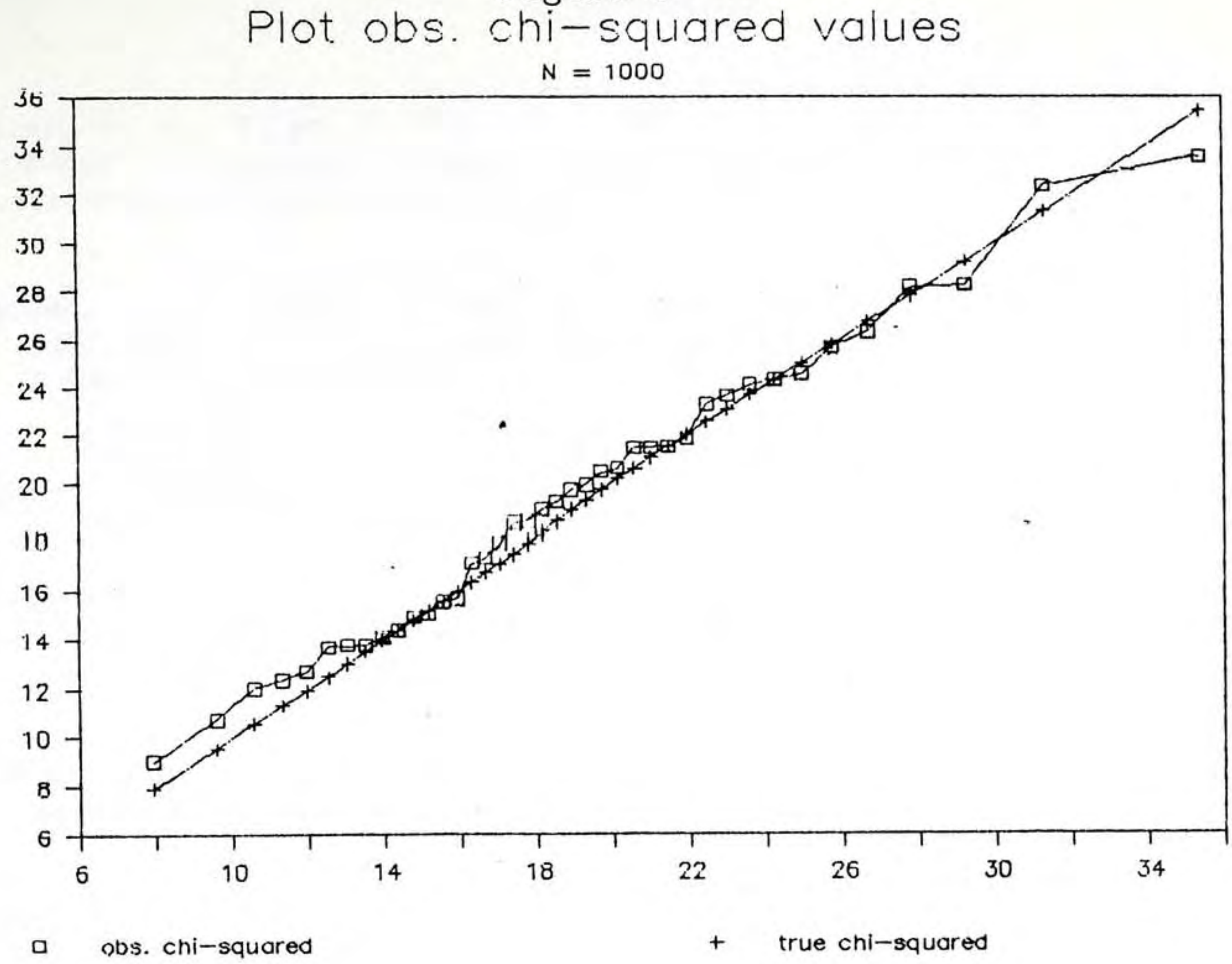


Figure 5


Plot obs. chi-squared values
N = 1000
36
34
32
30
28
26
24
22
20
18
16
14
12
10
8
6
6
10
14
18
22
26
30
34
obs. chi-squared
true chi-squared

References

Bentler, P. M., & Lee, S. Y. (1978). Matrix derivatives with chainrule and rules for simple, Hadamard, and Kronecker products. *Journal of Mathematical Psychology, 17*, 255-262.

Ferguson, T. W. (1958). A method of generating best asymptotically normal estimates with applications to the estimation of bacterial densities. *Annals of Mathematical Statistics, 29*, 1046-1062.

Gong, G. & Samaniego, F. J. (1981). Pseudo Maximum likelihood estimation : Theory and applications. *The Annals of Statistics, 9*, 861-869.

IMSL Library (1975). *International Mathematical and Statistical Libraries (ed. 5)*. Houston, Texas.

Johnson, N. L., & Kotz, S. (1972). *Distributions in statistics : Continuous multivariate distributions*. New York : Wiley.

Lawley, D. N., & Maxwell, A. E. (1971). *Factor Analysis as a Statistical Method*. London : Butterworhts.

Lee, S. Y., & Jennrich, R. I. (1979). A study of algorithms for covariance structure analysis with specific comparisons using factor analysis. *Psychometrika, 44*, 99-113.

Lee, S. Y., Poon, W. Y., & Bentler, P. M. (1990a). Full maximum likelihood analysis of structural equation models with polytomous variables. *Statistics and Probability letters, 9*, 91-97.

Lee, S. Y., Poon, W. Y., & Bentler, P. M. (1990b). A three-stage estimation procedure for structural equation models with polytomous variables. *Psychometrika, 55*, 45-51.

Olsson, U. (1979a). On the robustness of factor analysis against crude classification of the observations. *Multivariate Behavioral Research, 14*, 485-500.

Parke, W. R. (1986). Pseudo maximum likelihood estimation : the asymptotic distribution. *The Annals of Statistics, 9*, 355-357.

Pierce, D. A. (1982). The asymptotic effect of substituting estimators for parameters in certain types of statistics. *The Annals of Statistics, 10,* 475-478.

Poon, W. Y., & Lee, S. Y. (1987). Maximum likelihood estimation of multivariate polyserial and polychoric correlation coefficients. *Psychometrika*, *52*, 409-430.

Poon, W. Y., Lee, S. Y., & Bentler, P. M. (1991). Maximum likelihood estimation in a model with interval data. *Journal of Applied Statistics*, in press.

Poon, W. Y., & Lee, S. Y. (1991). Covariance structure analysis with interval data. Submitted for publication.